

TECNOLOGÍA COGNITIVA APLICADA A LA COMUNICACIÓN EN LAS AULAS
VIRTUALES

GUSTAVO IVAN ARCE MILLAN

OMAR DAVID SANCHEZ ORTIZ

UNIVERSIDAD PILOTO DE COLOMBIA
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA DE SISTEMAS
BOGOTÁ D.C.
2019

TECNOLOGÍA COGNITIVA APLICADA A LA COMUNICACIÓN EN LAS AULAS
VIRTUALES

GUSTAVO IVAN ARCE MILLAN
1310215

OMAR DAVID SANCHEZ ORTIZ
1311619

PROYECTO DE GRADO PARA OPTAR AL TÍTULO DE
INGENIERO DE SISTEMAS

ASESOR
GILBERTO PEDRAZA GARCÍA PhD

UNIVERSIDAD PILOTO DE COLOMBIA
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA DE SISTEMAS
BOGOTÁ D.C.
2019

Nota de aceptación.

Firma presidente del jurado

Firma coordinador del jurado

Firma calificadora del jurado

Bogotá D.C., _____

DEDICATORIA

A todos los profesores que aportaron en mi la semilla de la curiosidad y me han exigido autosuperarme a mi mismo cada dia

A mi madre, mi padre y mis dos hermanos que me enseñan algo distinto cada dia y a que vivir con felcicidad y satisfaccion es lo importante en la vida.

A mis jefes y mis compañeros de trabajo que me han ayudado a indagar en nuevas tecnologias y ser uno de los motores para comenzar este proyecto

AGRADECIMIENTOS

A nuestro asesor, el profesor Gilberto Pedraza Garcia, por ayudarnos en el proceso de elaboración de trabajo de grado, para hoy poder volver nuestra idea una realidad

TABLA DE CONTENIDO

	<i>Pág.</i>
1. DESCRIPCIÓN DEL PROBLEMA	16
1.1. PLANTEAMIENTO DEL PROBLEMA	16
1.2. ANÁLISIS DEL PROBLEMA	19
1.3. JUSTIFICACIÓN	20
1.4. OBJETIVOS	20
OBJETIVO GENERAL.	20
1.4.1 OBJETIVOS ESPECÍFICOS	20
1.5. LÍMITES Y ALCANCES	21
2. DISEÑO METODOLÓGICO	22
2.1. HIPOTESIS	22
2.2. VARIABLES INDEPENDIENTES	22
2.3. VARIABLES DEPENDIENTES	22
2.4. VARIABLES INTERVINIENTES	22
3. MARCO REFERENCIAL	23
3.1. MARCO TEÓRICO	23
3.1.1 CIENCIA Y TECNOLOGÍA	24
3.1.2 COGNICIÓN	26
3.1.3 INTELIGENCIA ARTIFICIAL	29
3.1.4 SCRUM	30
3.1.5 SOA	35
3.2. MARCO CONCEPTUAL	39
3.2.1 E-LEARNING	39
3.2.2 TECNOLOGIA COGNITIVA	41
3.3. MARCO DE ANTECEDENTES	48
3.3.1 HERRAMIENTAS DE CREACION	48
3.3.2 PRODUCTOS BASADOS EN CHATBOTS	52
4. PLAN DE DESARROLLO	54
4.1. IDENTIFICACIÓN DE INVOLUCRADOS	54
4.2. ALCANCE Y DEFINICIÓN DE TRABAJO	56
4.3. PLAN DE COMUNICACIÓN	57
4.4. CRONOGRAMA	58

5. DISEÑO DE PROTOTIPO	60
5.1. ARQUITECTURA DE SOFTWARE	66
5.2. ESTRUCTURA FISICA DE LA SOLUCION	68
5.2.1 Escenarios de Calidad	69
5.3. DIAGRAMA DE CLASES	74
5.4. INTERACCIONES SISTEMA	75
5.4.1 INTERACCION ESTUDIANTE-AULA VIRTUAL	75
5.4.2 CALIFICACIÓN RESPUESTAS	76
5.4.3 VISUALIZACIÓN PREGUNTAS Y RESPUESTAS	76
5.5. BASE DE DATOS	77
5.5.1 DIAGRAMA DE BASE DE DATOS	77
5.5.2 SCRIPTS	78
6. IMPLEMENTACION	80
6.1. ESPECIFICACIONES TÉCNICAS	80
6.2. FRONT-END	80
6.3. BACK-END	80
6.4. WATSON DISCOVERY	80
6.5. DIAGRAMAS FRONT-END	81
6.6. DIAGRAMAS FRONT END	83
7. ACEPTACIÓN DE USUARIOS	85
8. REQUERIMIENTO DE VALOR AGREGADO	90
8.1. WATSON DISCOVERY SERVICE	90
8.1.1 CONFIGURACIÓN DE DOCUMENTOS	91
8.1.2 ENRIQUECER DOCUMENTOS	93
8.1.3 ENTRENAR LA HERRAMIENTA PARA MEJORAR LOS RESULTADOS	94
9. RECOMENDACIONES	96
9.1. INTEGRACION CON LA PLATAFORMA E-LEARNING	96
9.2. ENTRADA/SALIDA POR VOZ	96
9.3. PREGUNTAS FRECUENTES	96
10. CONCLUSIONES	97
10.1. APORTACIONES DE LA TESIS	97
BIBLIOGRAFIA	98
ANEXOS	101

ANEXO A. FORMATO DE ENCUESTA	101
ANEXO B. PLAN DE PRUEBAS	105
• Alcance	105
• Fuera del alcance	105
• Estrategia	106
• Equipo de trabajo	106
• Prerrequisitos	107
• Supuestos	107
• Limitaciones	107
ANEXO C. Casos de prueba	108
10.1.1 Respuesta del chat	108
10.1.2 Calificación de la interacción	109
10.1.3 Calificación de la interacción	110
10.1.4 Visualización de graficas	112
ANEXO D. SQL vs NoSQL	114
• Cuando usar SQL	115
• Cuando usar no SQL	115
ANEXO E. DIAGRAMA INDISPONIBILIDAD HERRAMIENTAS SÍNCRONAS	116
ANEXO F. COSTOS Y LICENCIAMIENTO	117
• Alojamiento de la aplicación	117
• Servicio Watson Discovery (Malaeb, 2017)	118
ANEXO G. Base de datos	119
ANEXO H. PRUEBAS DE LA HERRAMIENTA	120
• Precisión y Recall de K	120
• Lo Recomendado con lo relevante	121
• Resultados	123

LISTA DE FIGURAS

	Pag.
Figura 1. Preferencia de uso herramientas sincrónicas vs. Asincrónicas.	17
Figura 2. Desventajas del uso del Skype	17
Figura 3. Desventajas del uso del Chat en línea.	18
Figura 4. Ciclo de desarrollo ágil..	33
Figura 5. Valor aportado por SOA	35
Figura 6. SOA desde el punto de vista de la tecnología.	36
Figura 7. Agilidad en el negocio articulada por SOA	38
Figura 8. E-Learning sincrónico y asincrónico.	40
Figura 9. características Watson Discovery.	43
Figura 10. Aprendizaje no supervisado	47
Figura 11. Aprendizaje reforzado	48
Figura 12. Intenciones en DialogFlow	49
Figura 13. Diálogos en DialogFlow	50
Figura 14. Ej. Estructura de Implementación de Watson Assistant	51
Figura 15. Ejemplo EduarBot	52
Figura 16. Jerarquía de involucrados (Autores).	54
Figura 17. Trello (Autores).	59
Figura 18. Vista de contexto (Autores)	66
Figura 19. Arquitectura de Software (Autores))	67
Figura 20. Diagrama de Componentes Solución del Problema (Autores).	68
Figura 21. Despliegue (nivel de instancia-Autores)	69
Figura 22. Diagrama de Clases (Autores).	74
Figura 23. Interacción estudiante Aula virtual (Autores).	75
Figura 24. Diagrama calificación respuestas (Autores).	76
Figura 25. Diagrama visualización preguntas y respuestas (Autores).	76
Figura 26. Diagrama base de datos. (Autores).	77
Figura 27. Interfaz Watson Discovery (Autores).	81
Figura 28. Diagrama Front-End Chat sin desplegar (Autores).	81
Figura 29. Diagrama Front-End Chat Desplegado (Autores).	82
Figura 30 Diagrama Front-End Respuesta del Chat (Autores).	82
Figura 31. Diagrama Front-End Calificación (Autores).	83
Figura 32. Diagrama Back-End Interacción (Autores).	83
Figura 33. Diagrama Back-End Calificación (Autores).	84
Figura 34 Diagrama Back-End Visualización (Autores).	84
Figura 35. Claridad en la información de los cursos virtuales.	85
Figura 36. Opciones utilizadas por los estudiantes para aclarar temas de las aulas virtuales.	85
Figura 37. Medios de comunicación más usados para contactar al docente.	86

Figura 38. Tiempo de respuesta de los instructores (Autores).	87
Figura 39. Relevancia de la información otorgada por Arquimides (Autores)	87
Figura 40. Tiempo de respuesta de Arquimides. (Autores)	88
Figura 41. Utilidad de Arquímedes en las aulas virtuales (Autores)	88
Figura 42. Interés de Arquimides en las aulas virtuales (Autores)	89
Ilustración 43. Indisponibilidad (Autores)	116
Figura 44. Planes de ClearDB	119

LISTA DE TABLAS

	Pag.
Tabla 1. Involucrados del proyecto (Autores).	54
Tabla 2. Definición de actividades (Autores).	56
Tabla 3. Plan de comunicación (Autores).	57
Tabla 4. Cronograma (Autores).	58
Tabla 5. Caso CH1 (Autores)	108
Tabla 6. Caso CH1 (Autores)	109
Tabla 7. Caso CH1 (Autores)	110
Tabla 8. Caso CH1 (Autores)	112
Tabla 9. SQL vs NoSQL (Autores)	114
Tabla 10. Comparación cuentas IBM Cloud	117
Tabla 11. Tipos de plan en Watson Discovery	118
Tabla 12. Resultados del Precision@K	123

GLOSARIO

ARTIFICIAL. herramientas diseñadas por el hombre para imitar o mejorar características de la naturaleza.

ASINCRÓNO. procesos, funciones o acciones independientes que se ejecutan en diferentes tiempos sin depender unas de otras.

AUTOMATIZACIÓN. proceso mediante el cual se convierten procesos manuales en automáticos a través de la tecnología, sin necesidad de intervención por parte de un humano.

BUSQUEDA ENRIQUECIDA. hace referencia a las búsquedas de información determinadas por ciertos lineamientos o criterios que le permiten encontrar de una forma más rápida y acertada lo que se desea.

CANAL DE COMUNICACIÓN. medio por el cual dos o más personas pueden interactuar y expresar sus opiniones o puntos de vista.

CHATBOT. herramienta tecnología en forma de chat, que permite la interacción entre un humano y una máquina.

CLOUD COMPUTING. en español computación en la nube, es un espacio en internet el cual permite almacenar información, aplicaciones o servicios.

COGNITIVO. es un adjetivo que se relaciona con la capacidad de aprender, modificar acciones por la experiencia.

DATOS ESTRUCTURADOS. son aquellos datos que están diseñados para que los lenguajes de programación entiendan su forma de ser.

DATOS NO ESTRUCTURADOS. son aquellos datos que no permiten identificar su forma interna, impidiendo así extraer información relevante de estos.

E-LEARNING. son plataformas virtuales que sirven para desarrollar o complementar cursos de aprendizaje apoyándose en herramientas didácticas de tipo tecnológico.

HERRAMIENTA COGNITIVA. con la combinación de tecnología cognitiva que incluyen algoritmos de aprendizaje automático y prácticas de procesamiento de información se crean herramientas cognitivas, que buscan solventar una rama en específico, como la capacidad de entender la voz o relacionar información.

NORMALIZACION DE DATOS. es la acción que permite convertir los datos no estructurados en datos estructurados, que permitan convertir estos en información valiosa y relevante.

NPL. proceso por el cual se estudia las interacciones entre las máquinas y el lenguaje natural de los humanos.

SAAS. acrónimo de “Software as a Service” y se refiere a una forma de distribución de software en el cual se expone la funcionalidad del software con la infraestructura y arquitectura de la solución transparentes para el cliente.

SINCRÓNO. procesos, funciones o acciones independientes que se ejecutan de forma secuencial en el tiempo.

SOA. arquitectura Orientada a Servicios, se refiere a los tipos de arquitecturas que se basan en exponer módulos de funcionalidad que reciben información, procesan y retornan otra que normalmente consume otro componente de la cadena lógica.

TECNOLOGÍA COGNITIVA. son las tecnologías que buscan funcionar como las actividades cerebrales humanas como comprender, escuchar o hablar. El procesamiento de lenguaje natural, el reconocimiento de patrones y la síntesis de voz son las ramas de la tecnología cognitiva más populares.

RESUMEN

La educación en su proceso de mejorar y adaptarse a la era tecnológica empezó a explorar nuevas soluciones apoyándose para esto de los canales virtuales, lo que cambió sustancialmente la comunicación de los usuarios ya que estos prefieren utilizar canales en tiempo real (síncronos) pero son llevados a optar por los canales asíncronos porque les permite una comunicación indiferente de la disponibilidad de los interlocutores.

En la educación tradicional la comunicación entre personas es a través de un canal sincrónico, con el fin de implementar la comunicación sincrónica se analizaron las diferentes soluciones tecnológicas que permitieran esta adaptación, optando por la tecnología cognitiva ya que se basa en la interacción entre humano y tecnología.

De acuerdo con lo anterior, se hace necesario el desarrollo de un software que sirva como canal sincrónico para suplir la disponibilidad de los actores involucrados y así permitir la comunicación preferida por los usuarios.

Palabras Claves. Tecnologías Cognitivas, Aulas virtuales, Desarrollo de Software, Comunicación Sincrónica.

INTRODUCCIÓN

La accesibilidad a la educación como uno de los aportes de la tecnología se refleja en la implementación y desarrollo de aulas virtuales los cuales también son elaborados con el fin de mejorar la calidad de la enseñanza en los estudiantes a través de diferentes técnicas o metodologías de aprendizaje, algunas de estas son los materiales didácticos y los espacios que se crean para generar un canal de comunicación entre estudiantes y docentes.

La elaboración de este proyecto busca implementar la tecnología cognitiva en el campo de la educación y mejorar la comunicación sincrónica, así es como nace la idea de desarrollar un protio de software que supla la indisponibilidad de los interlocutores habilitando un nuevo canal sincrónico en las aulas virtuales.

1. DESCRIPCIÓN DEL PROBLEMA

1.1. PLANTEAMIENTO DEL PROBLEMA

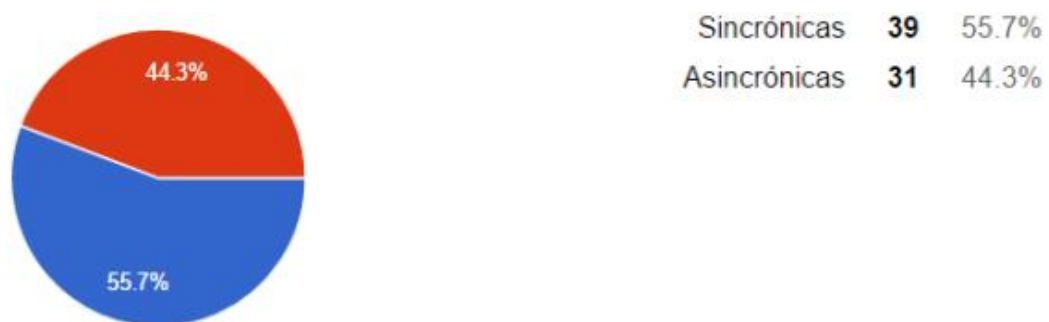
El "aula virtual" como herramienta educativa en la enseñanza a distancia, es el medio en el cual los educadores y educandos se encuentran para realizar actividades que conducen al aprendizaje (Scagnoli, 2001), es decir, es un sistema compuesto por actividades que permiten la adquisición, aplicación y evaluación de conocimientos referentes a temas o contenidos pensados para desarrollar un curso. Para llevar a cabo esto, el aula virtual cuenta con espacios para estas actividades de comunicación, algunas que no necesitan respuesta inmediata o destinatario concreto son llamadas asincrónicas como lo son foros de discusión y bibliotecas virtuales y otras que involucran interacción continua y se logran en tiempos cortos son llamadas sincrónicas como video llamadas y chats. "La comunicación sincrónica es el intercambio de información por Internet en tiempo real y la comunicación asincrónica se establece entre dos o más personas de manera diferida en el tiempo, es decir, cuando no existe coincidencia temporal"¹.

La comunicación sincrónica no se logra la mayoría de las veces ya que es necesario de la disponibilidad del profesor y los estudiantes,. Según el estudio realizado por Páez Edwin cuyo "objetivo de la investigación fue evaluar el uso de herramientas sincrónicas y asincrónicas en los procesos de aprendizaje a nivel de entornos virtuales en seis cursos de las ciencias agropecuarias de la Universidad Nacional Abierta y a Distancia – UNAD"² conformada por un total de 360 estudiantes, y una muestra de 70 de ellos, se concluyó que las herramientas de tipo asincrónico son más utilizadas por los estudiantes, sin embargo estos prefieren usar las herramientas sincrónicas las cuales poseen algunas limitantes lo que hace que su uso dentro de las clases virtuales se vea disminuido.

¹ LAMI Rodríguez del Rey, L. E., PEREZ Fleites, M. G., & Rodríguez del Rey Rodríguez, M. E. (2016). Las herramientas de comunicación sincrónica y asincrónica en la clase presencial. Revista Conrado [seriada en línea], 12 (56), pp. 84-89. Recuperado de <http://conrado.ucf.edu.cu/>

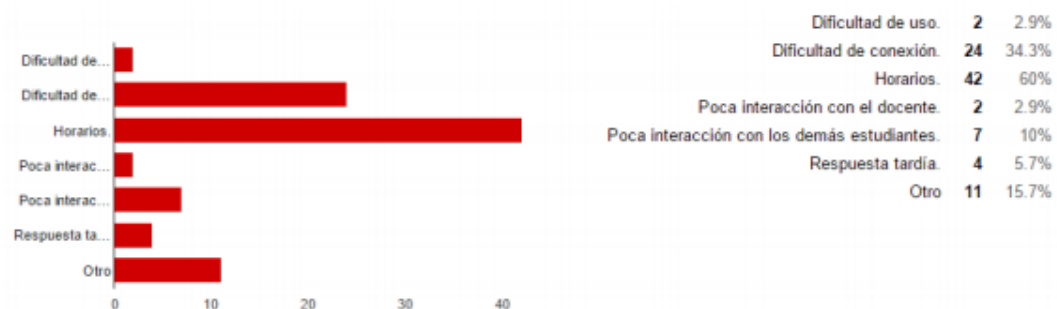
² PAEZ, Edwin. CORREDOR, Emma. FONSECA Jorge. Revista Ciencia y Agricultura Vol.13. Evaluación del uso de herramientas sincrónicas y asincrónicas en procesos de formación de las ciencias agropecuarias. Tunja (Boyacá). [en línea] 2016. P.79. 20 abril de 2018. Disponible en <https://dialnet.unirioja.es/descarga/articulo/5560523.pdf>

Figura 1. Preferencia de uso herramientas sincrónicas vs. Asincrónicas³.



Las limitantes a las que se refiere el autor se pueden evidenciar cuando este menciona las desventajas de la comunicación sincrónica como lo es Skype, la cual permite una comunicación en tiempo real con el docente pero disminuye su uso al necesitar de horarios de disponibilidad entre ambas partes, respecto a las desventajas del uso del Skype, 42 estudiantes (correspondientes al 60 % del total) reconocen que la principal limitante son los horarios, puesto que en los cursos se brinda un horario específico para desarrollar los procesos de comunicación con Skype⁴.

Figura 2. Desventajas del uso del Skype⁵



³ Ibíd., p.82.

⁴ Ibíd., p.86.

⁵ Ibíd., p.86.

Otra de las limitantes que se pueden observar en el estudio en cuanto a las herramientas sincrónicas como lo es un chat en línea se evidencia cuando 38 estudiantes (correspondientes al 54.3 % del total) reconocen que la principal limitación está relacionada con los horarios, ya que al tratarse de una herramienta sincrónica exige el establecimiento de unos horarios de atención específicos, que en algunas ocasiones coinciden con otras actividades laborales o personales de los estudiantes.⁶

Figura 3. Desventajas del uso del Chat en línea⁷.



Según lo visto anteriormente se hace necesario desarrollar una herramienta que supla la indisponibilidad en la comunicación sincrónica entre el docente y el estudiante permitiendo la solución de dudas dentro del aula virtual en tiempo real y sin las limitantes de las herramientas utilizadas convencionalmente.

Las aulas virtuales en la universidad son comúnmente utilizadas en cursos presenciales que apoyan el aprendizaje autónomo, el cual cuenta con el 75% del esfuerzo académico del estudiante (una clase presencial son 3 horas de aprendizaje autónomo). Al tener un buen porcentaje en el aprendizaje del estudiante es importante poder reforzar aspectos que puedan mejorar el aprendizaje autónomo como la disponibilidad.

Los temas de fundamentos de programación son la base de las carreras de ingeniería de sistemas y afines, que abarcan temas que tienen una base muy fuerte en conceptos específicos que son difíciles de entender sin aclaraciones específicas de estos, y viendo las cifras de la investigación es difícil obtener esta información con los canales actuales de comunicación.

⁶ *Ibíd.*, p.88.

⁷ *Ibíd.*, p.88.

1.2. ANÁLISIS DEL PROBLEMA

Cómo se pudo evidenciar en las limitantes de los canales sincrónicos como Skype la disponibilidad es un factor crítico en este tipo de comunicación, por lo que sería la primera variable a resolver, esta comunicación necesita la disponibilidad de dos interlocutores (en este caso el profesor y alumno) para hacerse efectiva, normalmente en un curso muchos estudiantes tienen un profesor que les asesora, por lo que la disponibilidad del profesor tiene que ser repartida por la cantidad de estudiantes que toman el curso, haciéndola más escasa.

Como punto de mejora la disponibilidad del profesor es la que se tiene que mejorar, este es necesitado por el conocimiento y experiencias en los temas del curso.

Con el planteamiento anterior se indagó sobre las posibles soluciones viables para mejorar la comunicación del estudiante con el profesor en el curso virtual, de las cuales se plantearon tres soluciones.

- Apoyar el curso con otro docente el cual mejoraría la comunicación sincrónica con los estudiantes al brindar más espacios de comunicación.
- Habilitar canales informales como pueden ser las redes sociales para la comunicación entre docentes y estudiantes.
- Crear un nuevo canal de comunicación sincrónico que se alimente de documentación de la clase y esté respaldado bajo supervisión del docente asegurando que las soluciones de las dudas a los estudiantes mejoren con la interacción de estos.

Al analizar las soluciones propuestas, se determinó que la primera implicaría una inversión en un recurso adicional con disponibilidad exclusiva para atender un canal de comunicación con el estudiante, la segunda opción implicaría invadir los espacios privados y personales de los docentes, en cambio, la tercera opción genera un costo inicial de implementación y permite habilitar un canal de comunicación en disposición total para aclarar las dudas de los estudiantes referentes al curso.

Es así como nace la idea de desarrollar una herramienta que permita una comunicación síncrona que no se limite por la disponibilidad del docente, y a su vez permita que este se apoye en la herramienta para mejorar la calidad de aprendizaje de los estudiantes por medio de la supervisión y entrenamiento continuo de esta.

Esta herramienta se apoyará en las tecnologías cognitivas las cuales se han impuesto en los mercados actuales y han dado apoyo a sectores de la salud y de la banca, dando a conocer sus beneficios en el análisis y transformación de las grandes cantidades de datos, por lo que este sería un acercamiento al sector de la educación y una base para la implementación en otras asignaturas.

1.3. JUSTIFICACIÓN

Debido al problema expuesto anteriormente, es necesario que la disponibilidad de los canales síncronos de las aulas virtuales sea mejorada como uno de los factores primordiales en la generación de contenido de apoyo, ya que los canales síncronos son los de mayor preferencia para los estudiantes y con mejor enfoque a solucionar dudas puntuales de estos, mejorando la disponibilidad de la comunicación sincrónica se beneficia a los estudiantes que toman el curso al poder preguntar de los temas en el aula virtual de manera rápida donde las respuestas están acotadas a los temas de la asignatura y alineadas con la información del curso.

El proyecto debe buscar la inclusión de tecnologías que aporten al proceso de clasificar información y enriquecer la búsqueda de contenidos, para facilitar respuestas alineadas a las temáticas, esto con el objetivo de optimizar el tiempo de los profesores en estas, para que su conocimiento en los contenidos pueda ser transmitido más eficientemente sin afectar los tiempos de esfuerzo. Este proyecto debe plasmar los conocimientos obtenidos en la carrera relacionándolo con las investigaciones de nuevas tecnologías para un óptimo desarrollo.

1.4. OBJETIVOS

OBJETIVO GENERAL.

Solucionar la indisponibilidad relacionada a los canales de comunicación sincrónico en las aulas virtuales.

1.4.1 OBJETIVOS ESPECÍFICOS

- Analizar el estado del problema teniendo en cuenta el estado actual del mercado que se involucra.

- Crear un canal de tipo sincrónico que supla la disponibilidad del profesor o experto en tema en el proceso de aprendizaje del alumno.
- Validar que el canal transmita el conocimiento del profesor o experto, para que el curso ofrezca la información correcta a los estudiantes.

1.5. LÍMITES Y ALCANCES

El desarrollo del prototipo de software comprenderá la integración de una herramienta que permita la estructuración de los datos que estén en lenguaje natural, para así poder extraer la información relevante para los estudiantes y responder las preguntas que puedan surgir de los temas del curso, contará únicamente con el entrenamiento de información sobre Fundamentos de programación. Esta herramienta tiene un costo que varía dependiendo de las consultas (\$0.115 USD por consulta nueva) y por documentos subidos por hora (\$0.01599 USD), la herramienta brinda los primeros 2000 documentos y 1000 consultas sin costo.

Este software contará con una interfaz de chat sencilla que permita la interacción de la información no estructurada con el estudiante, adicionalmente, esta herramienta contará con tecnología cognitiva por lo que requerirá un proceso de aprendizaje supervisado la cual contendrá una base de datos y la visualización de una/dos gráficas. La aplicación depende de una infraestructura y un servidor de aplicaciones (servidor Tomcat para pruebas) que tiene un costo variable según el proveedor, en el caso del prototipo se utilizará el ambiente IBM Cloud con 256 MB de memoria el cual no tiene costo. La aplicación requiere de una base de datos relacional que varía dependiendo de la capacidad y tráfico de la aplicación, para el prototipo se utilizará una instancia ClearDB basada en MYSQL capacidad de 10MB y 5 conexiones simultaneas.

El diseño y desarrollo de software será un prototipo únicamente para los temas de fundamentos de programación, con una formación de datos inicial basado en la documentación que se considera relevante para el curso, este prototipo no siempre contestará correctamente ya que esto depende del seguimiento del bot con las herramientas de entrenamiento.

El enfoque del software es interactivo y cognitivo, por lo que contará de una interfaz sencilla que aún no será implementada en el Aulas Virtuales como Moodle, dejando la posibilidad en un futuro de ser acoplada en un curso.

El proyecto está planeado en ejecutarse dividido en 6 Sprint de uno o dos semanas, con un equipo conformado por dos desarrolladores.

El proyecto contará con un análisis basado en el problema, buscando medir en variables de tiempo y adopción la herramienta con respecto a los canales sincrónicos tradicionales, en se dará respuesta a la validación o no de la hipótesis.

2. DISEÑO METODOLÓGICO

2.1. HIPOTESIS

- la tecnología cognitiva permite implementar canales de comunicación sincrónica en las aulas virtuales sin depender de la disponibilidad de los interlocutores.

2.2. VARIABLES INDEPENDIENTES

- Canal de comunicación sincrónica.

2.3. VARIABLES DEPENDIENTES

- Aulas virtuales
- Disponibilidad del conocimiento.

2.4. VARIABLES INTERVINIENTES

- Proceso de aprendizaje.
- Experticia del docente.

3. MARCO REFERENCIAL

3.1. MARCO TEÓRICO

“El país debe continuar mejorando la calidad y pertinencia de la educación superior y a la vez atraer y acoger un alumnado cada vez más cuantioso y diverso. Esto implicará mayores esfuerzos para orientar y apoyar a los estudiantes a lo largo del desarrollo de su educación superior, especialmente los jóvenes de orígenes menos favorecidos, y abordar los obstáculos económicos y geográficos para el acceso”⁸.

Según la revisión detallada de la política de educación superior en Colombia realizada por la Organización para la Cooperación y el Desarrollo Económicos (OCDE) se identificaron varios logros y problemas, uno de estos a solucionar es la calidad, para esto es necesario proporcionar una educación consistente en el contexto nacional y regional tomando conciencia de la necesidad de eliminar las barreras de acceso a los estudiantes de regiones y grupos desfavorecidos, lo que implica permitir una mayor adquisición de competencias a través de programas pertinentes.

Con lo anterior, también se busca mejorar la retención de los estudiantes garantizando la finalización de sus estudios, lo que resalta la importancia del desarrollo de estrategias y ajustes de enfoques tradicionales de enseñanza.

Algunos enfoques favorables implementados por las instituciones de educación superior son.

- métodos de enseñanza enfocados en los estudiantes, menos dirigidos y más participativos.
- formar profesores para diagnosticar las áreas de debilidad de los estudiantes y ofrecerles ayuda adaptada.

⁸ Organización para la Cooperación y el Desarrollo Económicos (OCDE). Ministerio de Educación Nacional. Educación en Colombia. Educación superior en Colombia. 2016, p.268.

En cuanto a mejorar la cobertura en las regiones se opta por el aprendizaje a distancia.

Según la encuesta realizada por la OCDE sobre el aprendizaje en línea en América latina, se evidencio el intento por parte de las entidades de educación superior el tratar de ampliar sus programas a través del uso de herramientas en internet y del aprendizaje en línea El 74% de las universidades de América Latina encuestadas tiene estrategias establecidas de aprendizaje en línea, y el 21% está desarrollando alguna de estas estrategias en la actualidad (OCDE, 2015c)⁹.

Estas entidades de educación superior establecidas en Colombia han optado por ofrecer programas enfocados en la educación virtual y a distancia, con el fin de otorgar acceso sin importar la dispersión regional de la población. Sin embargo, el potencial de la educación a distancia sigue siendo limitado ya que, si bien el registro de los estudiantes en cursos virtuales ha aumentado en comparación con los cursos presenciales, la tasa de participación sigue siendo baja, la OCDE sugiere que “La ampliación de estos programas, especialmente el aprendizaje a distancia en línea, podría ser una contribución importante para remediar los desbalances regionales en el futuro”¹⁰.

Algunos beneficios del aprendizaje a distancia incluyen la posibilidad infinita de ampliación y el costo de prestación relativamente bajo, una vez que los programas y los materiales del programa han sido desarrollados. Un aprendizaje combinado, en el cual los estudiantes aprenden de forma remota pero también se benefician de instrucción presencial, puede ser especialmente eficaz en situaciones como la de Colombia, donde los estudiantes pueden carecer de competencias de autonomía en el estudio.

3.1.1 CIENCIA Y TECNOLOGÍA

Para hablar de tecnología siempre se ha debido relación el concepto de ciencia, la cual se puede definir desde diferentes puntos de vistas según las formas de estudio y corrientes teóricas, estableciendo así a la ciencia como un área de estudio

3.1.1.1 CIENCIA

⁹Ibíd., p.302.

¹⁰ Ibíd., p.302

Como cita Pérez Eduardo en su texto Ciencia y tecnología al alcance de todos. John Ziman en su texto el conocimiento público analiza cuatro definiciones “la primera manifiesta que la ciencia es el dominio del medio que rodea al hombre”¹¹, esto quiere decir que la ciencia se enfoca en las aplicaciones del conocimiento científico y da prioridad a las cosas más que a las ideas.

En la segunda definición dice que “la ciencia es el estudio del mundo material”¹², considera que esta interpretación es aceptada en el pensamiento popular, quedando excluidas la ciencias exactas y sociales ya que se enfoca solo en lo material.

La tercera definición “la ciencia es el método experimental”¹³ donde se considera la ciencia como una metodología asociada a las pruebas y la observación, metodología que no todas las ciencias aplican como las matemáticas.

Y la última definición “la ciencia llega a la verdad por inferencias lógicas de observaciones empíricas”¹⁴, es decir, por medio de la observación de un fenómeno que no varía en su comportamiento este se define como un hecho básico o ley convirtiéndose así en una verdad parcial en incompleta ya que esta forma de investigación no se somete a errores ni contradicciones (principio de inducción).

La ciencia se puede definir desde diferentes puntos de vista, “la mayoría de los científicos y la mayor parte del público adoptan una u otra de las actitudes que se han señalado de acuerdo con su grado de preparación intelectual”¹⁵.

La ciencia al ser parte de la construcción social, su definición va más allá de la investigación y la metodología, es necesario involucrar en su dinámica a la sociedad como prioridad ya que estas investigaciones deben responder a las necesidades e intereses de todos¹⁶.

3.1.1.2 TECNOLOGÍA

¹¹ PEREZ Cevallos. Alex Eduardo. Ciencia y tecnología al alcance de todos. 2007, p.22.

¹² Ibíd., p.22.

¹³ Ibíd., p.22.

¹⁴ Ibíd., p.23.

¹⁵ Ibíd., p.23.

¹⁶ Ibíd., p.25.

De acuerdo con Aristóteles la tecnología “se relaciona de manera esencial con cosas y objetos creados artificialmente”¹⁷, es decir, la tecnología es el resultado del ingenio e inventiva del ser humano, son todos aquellos objetos o herramientas diseñadas artificialmente para facilitar las labores o necesidades de las personas, la época actual se puede catalogar como tecnológica ya que la mayoría de los seres humanos están altamente influenciados por la tecnología y en constatación de interacción con esta.

“Otra definición de tecnología dada por los intelectualistas representados por los académicos es considerar a la tecnología como ciencia aplicada”¹⁸. esto quiere decir que la tecnología surge del conocimiento adquirido a través de la práctica, aunque han surgido debates que afirman que la tecnología es una derivación de la ciencia y sin esta no existiría, podemos observar en la actualidad la creación de herramientas tecnológicas avanzadas que permiten un mejor estudio y/o descubrimientos, aportando nuevos métodos de investigación y desarrollo de nuevos conocimientos científicos.

3.1.2 COGNICIÓN

La cognición del latín cognoscere que se traduce como conocer y de allí se deriva el conocimiento el cual se alcanza mediante el ejercicio de las facultades mentales, afirmando la habilidad o la capacidad de un individuo para procesar la información a partir de la percepción. Consiste en la capacidad de aprender a través de las experiencias, de razonar a través de la observación para así formular solución a problemas encontrados, tomando decisiones basados en los procesos anteriormente descritos los cuales son mayormente características de los seres humanos.

3.1.2.1 MENTE

Es el conjunto de facultades cognitivas, que engloban procesos como la percepción, el pensamiento, la conciencia y la memoria, algunas corresponden a características del humano. “La mente no es únicamente el asiento de la parte cognitiva (pensamiento, percepción, memoria, inteligencia), sino también de la

¹⁷ Ibíd., p.26.

¹⁸ Ibíd., p.26.

parte afectiva”¹⁹, Se puede decir que la mente hace referencia al resultado de los procesos ejecutados por el cerebro, pero no solo su ejecución sino su interacción, profundización y correlación, siendo así capaz de crear y regular sentimientos, creencias y emociones en un ámbito no solo biológico, sino cultural y social.

3.1.2.2 RAZONAMIENTO

“El razonamiento es una relación entre juicios, no es verdadero ni falso”²⁰, es decir, el razonamiento, es la manera en la cual concluimos a través de la relación entre procesos de experiencia y conocimientos obtenidos el resultado a un determinado escenario o problema, generando así nueva información basado según lo anteriormente descrito en premisas obtenidas.

Existe cinco tipos de razonamientos.

Análisis. “El análisis consiste en separar los elementos de un todo o de objeto para estudiar más fácilmente”²¹. Es decir, partes el problema desde lo menos complejo a lo más complejo para así entender cada parte y luego su total.

Síntesis. “La síntesis en su sentido más amplio es la unificación de las partes en un todo”²². Es decir, la organización y la relación de partes de un problema u objeto complejo.

Analogía. “La analogía es un razonamiento en que se buscan semejanzas genéricas para proponerlas en el estudio de diferentes situaciones”²³. Este hace

¹⁹ MORIELLO Sergio. Cerebro, Mente, Cuerpo y Entorno. [en línea]. 24 febrero de 2018. Disponible en http://www.pensamientocomplejo.com.ar/docs/files/Moriello_Cerebro-mente-cuerpo-entorno.pdf

²⁰ WIECHERS Rivero. Lógica, versión icónica. Editorial humanismo y sentido. [en línea]. México 2009. 24 febrero de 2018. Disponible en http://cvonline.uaeh.edu.mx/Cursos/Bach_Virt/H0102/unidad_4/lec_r10315_el_razonamiento.pdf

²¹ Ibíd., p.2.

²² Ibíd., p.2.

referencia a la proporción es decir el razonamiento en objetos o sucesos que guardan cierta equivalencia.

Dedución. “Deducir significa pasar de lo general a lo particular”²⁴. Este concepto va de la mano con las proposiciones, es decir, si aceptamos una ley universal podemos acoger bajo estas varias situaciones que tengan el mismo comportamiento, pasando de lo general a lo menos general. El silogismo es la forma tradicional de deducción lógica, de la combinación de dos premisas que se relacionan, derivamos forzosamente una conclusión. Como ya dijimos, el Razonamiento Deductivo es una forma de conocimiento discursivo en el que pasamos del conocimiento general a otro general u otro menos general, a otro particular o singular. Un ejemplo en el que posamos de lo general a lo particular puede ser el siguiente.

- Todos los metales son buenos conductores de electricidad.
- La plata es un metal.
- Por lo tanto, la plata es buena conductora de metal.

Las dos primeras premisas están relacionadas. La palabra metal la relaciona, entonces podemos concluir algo sobre estas premisas, hay que aclarar que a la tercera oración no se llama premisa sino conclusión.

La diferencia entre el razonamiento deductivo y el razonamiento inductivo es el grado de certeza, un razonamiento deductivo si es correcto, no permite tener mayor grado de certeza en nuestras conclusiones y en el caso de inducción, si nuestro razonamiento es correcto, su nivel de certeza siempre será menor y en ocasiones casi nulo.

Inducción.

²³ Ibíd., p.3.

²⁴ Ibíd., p.3.

“La inducción es obtener principios explicativos generales a partir de los fenómenos o casos particulares”²⁵, Es decir a partir de la observación de fenómenos particulares y su razón de suceder, se pueden obtener nuevos fenómenos particulares que permiten generar nuevas teorías o leyes (principios generales).

- De lo particular a lo particular
- De lo particular a lo menos particular
- De lo particular a lo general.

Las conclusiones basadas en el razonamiento deductivo son altamente confiables, lo que nos garantiza un grado alto de certeza, a diferencia de esta, las conclusiones basadas en razonamiento de inducción son de baja confiabilidad.

3.1.3 INTELIGENCIA ARTIFICIAL

“La inteligencia artificial (IA) es una rama de las ciencias computacionales encargada de estudiar modelos de cómputo capaces de realizar actividades propias de los seres humanos con base a dos de sus características primordiales. el razonamiento y la conducta”²⁶.

Una de las capacidades que tiene el ser humano es la capacidad de entender y dar sentido a oraciones y textos de diferente forma variando del entorno o contexto tratado, aplicaciones de la inteligencia artificial en el procesamiento de lenguaje natural, esto habilita la posibilidad de explorar diferentes fuentes de información que antes eran complejas de entender, como libros y documentaciones, extraer información que la etiqúete en segmentos que puedan ser procesados a gran volumen.

²⁵ Ibíd., p.5.

²⁶ LOPEZ Bruno. Introducción a la inteligencia Artificial, Instituto tecnológico de Nuevo Laredo. [en línea]. 2007. 10 marzo 2018. Disponible en <http://www.itnuevolaredo.edu.mx/takeyas/Articulos/Inteligencia%20Artificial/ARTICULO%20Introduccion%20a%20la%20Inteligencia%20Artificial.pdf>

El objetivo de la inteligencia artificial es el de crear una máquina que se pueda hacer pasar por un ser humano, alguno de los principales ejemplos son el test de Turing, la máquina debe ser capaz de realizar procesamiento de lenguaje natural, razonamiento y aprendizaje.

3.1.4 SCRUM

Scrum es un marco de trabajo para el desarrollo de proyectos, es decir, una nueva forma de gestionar basados en la agilidad y flexibilidad que es llevada a cabo con el fin de crear un producto, servicio o resultado; Scrum sirve como marco de trabajo en el cual se pueden emplear varias técnicas y procesos.

El marco de trabajo Scrum consiste en equipos, roles, eventos, artefactos y reglas asociadas, gobernando las relaciones e interacciones entre ellos. “Se basa en la teoría del control de procesos empíricos asegurando el conocimiento de la experiencia para así tomar decisiones empleando un enfoque interactivo e incremental para optimizar la predictibilidad y el control del riesgo”²⁷.

Scrum es adecuado en las empresas en las que el desarrollo de los productos se realiza en entornos con.

- Incertidumbre.** proyectos en los cuales no se tiene un plan detallado del producto, generando autonomía y motivación en los equipos.
- Auto-organización.** equipos autónomos capaces de encontrar la estrategia para desarrollar la solución más adecuada, mejorando con la experiencia, y cohesivamente.
- Control.** se trata de trabajo en equipo regulado, donde se fomenta la creatividad y espontaneidad entre los colaboradores.

²⁷ SCHWABER Ken. La Guía Definitiva de Scrum. Las Reglas del Juego. [en línea]. 2016. 17 marzo de 2018. Disponible en <http://www.scrumguides.org/docs/scrumguide/v2016/2016-Scrum-Guide-Spanish.pdf>

•**Transmisión del conocimiento.** “Todo el mundo aprende de todo el mundo, los colaboradores participan en diferentes proyectos compartiendo su conocimiento a lo largo de la organización”²⁸.

3.1.4.1 ROLES

Personas comprometidas con el proceso y el proyecto Scrum.

Dueño del producto. Es la persona que toma las decisiones, es quien realmente conoce el negocio del cliente, es el responsable de maximizar el valor del producto y el trabajo del equipo de desarrollo, es el único que debe gestionar la lista del producto (Product Back log), es decir, se debe incluir las ideas del cliente materializadas en tareas y definiendo su respectiva prioridad.

Como se mencionó anteriormente, el dueño del producto es el único que puede modificar la lista del producto, por lo que cualquier cambio debe ser remitido a través de él y es él quien decide que cambios se deben realmente realizar, lo que genera una directriz tanto para el trabajo desarrollado por el equipo basado en los requerimientos expuestos en la lista, como para las tareas requeridas en base a las ideas de negocio que el cliente quiere ver reflejadas en el producto final.

Equipo de desarrollo. El equipo de desarrollo consiste en un grupo pequeño (entre 5-9 personas) de profesionales encargados del desarrollo del producto y el incremento de este se debe entregar al finalizar cada sprint, es decir, el avance del producto debe permitir su posible puesta en producción; para esto los equipos de desarrollo deben ser estructurados y empoderados generando autonomía, eficiencia y efectividad.

Además, el equipo de desarrollo debe ser auto organizado y multifuncional ya que bajo la metodología Scrum el equipo es quien debe encargarse de cumplir con los objetivos de la lista y el incremento del producto, aunque el equipo puede estar integrado con profesionales especializados en diferentes campos o dominios, la responsabilidad y el esfuerzo necesario para cumplir con las tareas y los objetivos recae en todo el grupo de desarrollo.

²⁸ TRIGAS Manuel. Gestión de Proyectos Informáticos. p.33. [en línea]. 17 marzo de 2018. Disponible en <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/17885/1/mtrigasTFC0612memoria.pdf>

Scrum Master. Es el responsable de asegurar que el trabajo del equipo se ajuste a las reglas de la metodología Scrum, también sirve como mediador entre las personas externas al equipo (gestores y cliente) para crear valor al trabajo desempeñado por el equipo.

El Scrum Master interactúa con el dueño del producto con el fin de establecer formas efectivas de gestionar la lista de producto, aclarar para el equipo los elementos incluidos en la lista ordenando estos por prioridad para maximizar su valor y facilitar los procesos de Scrum lo que asegura agilidad en el trabajo del equipo.

En cuanto a la función del Scrum Master con el equipo este debe guiarlos en un proceso de auto organización como se mencionó anteriormente, ayudar al equipo a generar un producto de alto valor, gestionar las interacciones fuera del equipo para que este se enfoque únicamente en los objetivos definidos y gestionar para realizar los procesos requeridos para la óptima implementación de la metodología Scrum.

Scrum al ser una metodología de desarrollo ágil tiene como base la creación de ciclos breves que se llaman Sprints.

Fases que definen el ciclo de desarrollo ágil.

- Concepto.** Definición de las características del producto y equipo encargado de su desarrollo.
- Especulación.** Se establecen límites del desarrollo como lo son tiempos y costos; se construye el producto a partir de ideas generales, esta fase es iterativa, revisando los requisitos generales, impacto, funcionalidad, fechas de la versión y esfuerzo estimado por cada iteración.
- Exploración.** Incremento del producto al añadir funcionalidades a la fase de especulación.
- Revisión.** El equipo realiza una comparación de lo construido con el objetivo final.
- Cierre.** Se entregará en la fecha acordada una versión del producto, esto no significa que el proyecto se haya finalizado, sino que pasará a una etapa de mejora o mantenimiento que hará que el producto se acerque al objetivo deseado.

Figura 4. Ciclo de desarrollo ágil.²⁹



3.1.4.2 BACK LOG

Es un documento que refleja los requisitos del sistema por prioridades, es aquí donde se planifica una lista de tareas para cada sprint. Para el seguimiento de estos Sprint cada tarea será evaluada en cuanto al trabajo realizado desde la última reunión, trabajo que se debe realizar antes de la siguiente reunión, inconvenientes surgió y soluciones posibles para continuar con lo planificado.

Scrum gestiona estas interacciones a través de reuniones diarias las cuales son uno de los elementos fundamentales de esta metodología denominadas **Daily**, estas pueden tener un tiempo de duración de 15 minutos para llevar a cabo la inspección del trabajo realizado diariamente haciendo un estimado del trabajo que se lograr completarse para el Sprint en el que se encuentre el equipo, permitiendo al final de Sprint entregar una versión del proyecto facilitando el feedback para el cliente.

En la revisión del Sprint requiere del equipo Scrum y los interesados para realizar una revisión de las tareas realizadas durante el Sprint y el avance en el producto

²⁹ Ibíd., p.34.

entregado, se trata de una reunión informal, no de seguimiento donde se presenta el incremento (es la parte producida del producto en el Sprint) fomentando la participación colaborativa del equipo y la retroalimentación.

La revisión del Sprint incluye los siguientes elementos.

- El dueño del producto expone los elementos de la lista de producto que se han terminado y cuáles no.
- El equipo menciona los puntos fuertes del trabajo realizado, los problemas encontrados y la forma en que se solucionaron.
- El equipo de desarrollo demuestra el trabajo terminado y responde preguntas acerca del incremento.
- El equipo realiza una realimentación, definiendo los pasos a seguir en los siguientes Sprints basados en la revisión del Sprint terminado.
- Revisión de la línea de tiempo, presupuesto y capacidades potenciales.

La retrospectiva del Sprint es una oportunidad de retroalimentación para el equipo según el trabajo realizado para así mejorar su productividad, esta retrospectiva es realizada entre la finalización de un sprint y el comienzo del siguiente.

Propósitos de la retrospectiva.

- Revisar el desempeño del equipo en cuanto a colaboración, comunicación y herramientas.
- Determinar según los elementos más importantes si estos se pudieron concretar o los inconvenientes que surgieron para poder mejorarlos.
- Crear según lo visto anteriormente planes de acción y retroalimentación para mejorar el desempeño del equipo Scrum.

El Scrum Master es quien lidera al equipo para diseñar procesos de mejora según lo realizado en el sprint, para esto el equipo debe realizar una retrospectiva crítica en la que se fomente la mejora en las prácticas de desarrollo y procesos de calidad con el fin de cumplir todos los objetivos de manera eficiente.

3.1.5 SOA

“Los sistemas informáticos tradicionales se han organizado en grandes bloques monolíticos que contienen tanto los procesos de negocio como sus funciones automatizadas. Estos sistemas han conseguido una gran mejora de productividad en las empresas, automatizando procesos de negocio, pero su concepción monolítica hace que los cambios y adaptaciones a las nuevas necesidades tiendan a ser más lentos y costosos de lo deseable”³⁰.

Con el fin de brindar una solución a lo anterior **SOA** busca generar agilidad a través de flexibilidad e integración de ecosistemas de tecnología “La arquitectura **SOA** establece un marco de diseño para la integración de aplicaciones independientes de manera que desde la red pueda accederse a sus funcionalidades, las cuales se ofrecen como servicios. La forma más habitual de implementarla es mediante Servicios Web, una tecnología basada en estándares e independiente de la plataforma, con la que SOA puede descomponer aplicaciones monolíticas en un conjunto de servicios e implementar esta funcionalidad en forma modular”³¹.

“La clave de la arquitectura SOA es “la abstracción de los procesos”, por la que los procesos de negocio se externalizan de las aplicaciones o soluciones y se exponen directamente al negocio. Esto permite que la ejecución, gestión, monitorización y modificación de dichos procesos puedan ser manejados directamente a nivel de negocio y de forma versátil, en vez de estar embebidos o imbricados en las aplicaciones.”³²

*Figura 5. Valor aportado por SOA*³³

³⁰ Centro de Alto rendimiento de Accenture (CAR) Arquitectura Orientada a Servicios (SOA). [en línea]. 2008. 29 abril de 2018. Disponible en http://www.kybele.etsii.urjc.es/docencia/IS_LADE/2012-2013/Material/CAR%20Accenture%20-%20SOA%5B1%5D.pdf

³¹ Microsoft Corporation. La Arquitectura Orientada a Servicios (SOA) de Microsoft aplicada al mundo real. 2006, p.3.

³² Centro de Alto rendimiento de Accenture, op. cit, p.6.

³³ *Ibíd.*, p.6.



Esta externalización de negocio se realiza a través de la exposición de servicios Web.

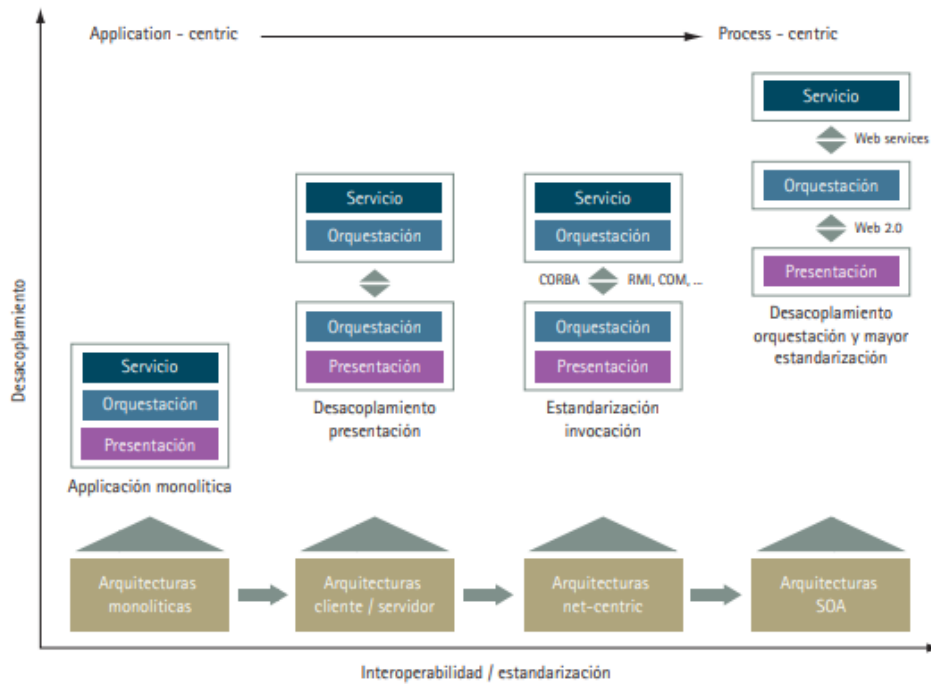
Los servicios Web son aplicaciones basadas en estándares para realizar el intercambio de información permitiendo la integración entre sistemas de cualquier plataforma (tanto de la organización como proveedores de servicios) y a su vez garantizar la comunicación entre estos, alguno de los estándar que hacen posible esta intercomunicación son "XML para la representación de datos, SOAP (Simple Object Access Protocol) para el intercambio de datos y el lenguaje WSDL (Web Services Description Language) para describir las funcionalidades de un servicio Web"³⁴.

*Figura 6. SOA desde el punto de vista de la tecnología*³⁵.

³⁴ Microsoft Corporation, op. cit, p.3.

³⁵ Ibíd., p.3.

Impacto de SOA en la evolución de las tecnologías de la información desde el punto de vista de desarrollo de aplicaciones

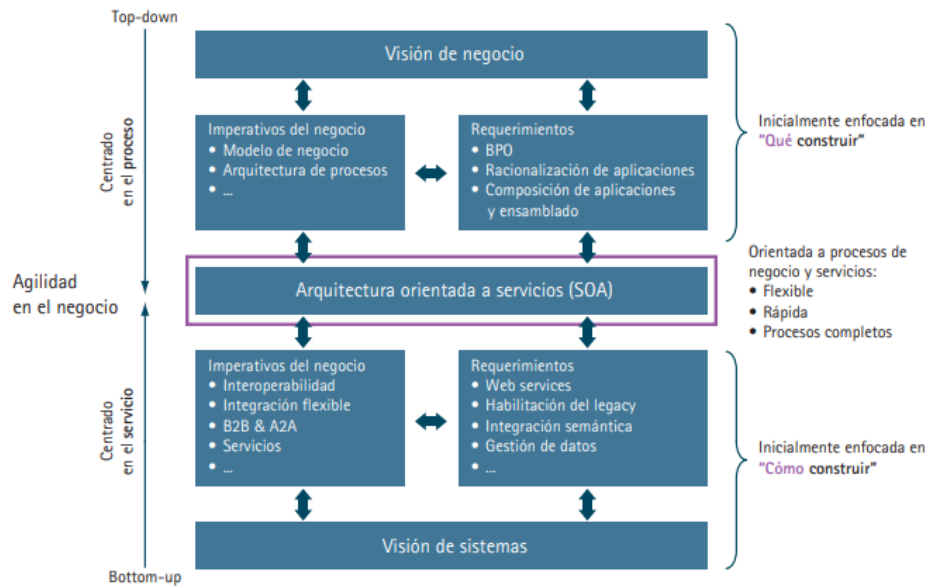


Algunos beneficios de SOA son.

- “La arquitectura SOA ayuda a mejorar la agilidad y flexibilidad de las organizaciones.
- La arquitectura SOA permite una “personalización masiva” de las tecnologías de la información.
- La arquitectura SOA permite la simplificación del desarrollo de soluciones mediante la utilización de estándares de la industria y capacidades comunes de industrialización.
- La arquitectura SOA permite aislar los sistemas frente a cambios generados por otras partes de la organización (protección de las inversiones realizadas).

- La arquitectura SOA permite alinear y acercar las áreas de tecnología y negocio³⁶.

Figura 7. Agilidad en el negocio articulada por SOA³⁷



³⁶ Ibíd., p.12.

³⁷ Ibíd., p.12.

3.2. MARCO CONCEPTUAL

3.2.1 E-LEARNING

“Las plataformas E-Learning, o plataformas virtuales, son cursos desarrollados en internet los cuales buscan dar soporte en la enseñanza y aprendizaje de los estudiantes que se integren a estos entornos, el mayor enfoque de estas plataformas es hacia los estudiante y cursos universitarios, comúnmente llamados aulas virtuales o clases virtuales que constituyen una realidad tecnológica permitiendo afirmar un fuerte cambio y transformación en los espacios de enseñanza tradicionales.”³⁸.

Los métodos que ofrece el E-Learning son enfocados en el estudio auto dirigido, es decir el estudiante debe buscar en estos cursos un apoyo para adquirir y profundizar conocimientos, a través de la retroalimentación asociada a actividades lúdicas que el curso le ofrezca, permitiéndole a este ajustarse a sus itinerarios y espacios disponibles para el aprendizaje.

Sin embargo, el E-Learning también está enfocado en apoyar a los instructores para que estos puedan optimizar el proceso de aprendizaje de sus estudiantes de una manera más efectiva y dinámica para lo cual “Existen dos enfoques globales para el e-Learning. el aprendizaje auto dirigido y el aprendizaje dirigido/facilitado por un instructor”³⁹.

³⁸ FERNANDEZ-PAMPILLON Ana. Universidad Complutense de Madrid. Las plataformas e-Learning para la enseñanza y el aprendizaje universitario en Internet. [en línea]. 17 febrero de 2018. Disponible en http://eprints.ucm.es/10682/1/capituloE_learning.pdf

³⁹ Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), Metodologías de E-Learning. GHIRARDINI Beatrice. [en línea]. 2014, p.11. 17 febrero de 2018. Disponible en http://www.fao.org/elearning/Sites/ELC/Docs/FAO_elearning_guide_es.pdf

3.2.1.1 COMPONENTES DEL E-LEARNING

- **Recursos de aprendizaje.** Son recursos no interactivos, como documentos, presentaciones, videos; entran en esta categoría ya que los estudiantes no realizan una interacción con ellos, es decir solo pueden leer o mirar el contenido.
- **Lecciones en línea.** El enfoque más común para el E-Learning auto dirigido se basa en la capacitación basada en la Web.
- **Material de apoyo.** Una lección es una secuencia lineal de páginas que pueden incluir textos, gráficos, animaciones, audio, video e interactividad en la forma de preguntas y comentarios. Las lecciones también pueden incluir bibliografía recomendada y enlaces a recursos en línea, así como información adicional sobre temas específicos”.

3.2.1.2 E-LEARNING SINCRÓNICO Y ASINCRÓNICO

“La característica más importante del e-Learning sincrónico son los eventos en tiempo real, eventos tales como comunicación entre dos personas o más personas como por ejemplo un chat; el e-Learning asincrónico se lleva a cabo independiente del tiempo un ejemplo de este puede ser el correo electrónico”⁴⁰.

*Figura 8. E-Learning sincrónico y asincrónico*⁴¹.

Sincrónico	Asincrónico	<p>La flexibilidad de la tecnología de Internet crea cierta ambigüedad en términos de lo que se considera sincrónico y asincrónico.</p> <p>Por ejemplo, las sesiones de video y audio pueden ser grabadas y puestas a disposición de los alumnos que no pueden asistir a clases.</p>
<ul style="list-style-type: none">> Chat e IM (mensajería instantánea)> Video y audio conferencias> Webcast en vivo> Intercambio de aplicaciones> Pizarra digital> Votaciones	<ul style="list-style-type: none">> Email> Foros de discusion> Wiki> Blog> Webcasting (transmisiones por web)	

⁴⁰ Ibíd., p.13.

⁴¹ Ibíd., p.13.

3.2.2 TECNOLOGIA COGNITIVA

La tecnología cognitiva, es la capacidad que tiene una máquina para recibir información, procesarla y obtener un conocimiento adicional mediante la experiencia (Cognición, como se había mencionado anteriormente...véase el numeral 2.1.2.), “la tecnología cognitiva es el conocimiento aplicado a las maquinas”⁴², esto quiere decir que las maquinas no solo recibirán información en base a su programación y elementos estructurados sino empezarán a crear nueva información a través de la experiencia basada en los datos del entorno. “Es decir, hoy estamos en la base de una plataforma donde las máquinas con sistemas de inteligencia artificial pueden tomar decisiones “inteligentes” basados en toda la información disponible en internet o sus bases de datos propias”⁴³.

La curva de aprendizaje en las maquinas aun esta por lo que se puede comparar como inteligencia blanda en los seres humanos y es el primer paso, en donde en base de una estructura se van trazando diferentes alternativas por medio de la información que el usuario brinda a sus respuestas, los aprendizajes supervisados son los que se pueden utilizar cuando se tiene una base de conocimientos que queremos que utilice como inicio, y que con este empiece a ofrecer información o clasificación sobre un tema, poblanamente esta primera iteración tenga información errada, en donde es necesario enseñarle en donde se ha equivocado para que trace cambios en sus respuestas

En un ejemplo se le muestra dos colores. amarillo y rojo en diferentes tonalidades, si le muestra un vino tinto posiblemente lo clasifique como rojo, pero si le muestras un naranja puede clasificarlo como amarillo o rojo por lo parecido en sus matices, si le muestras colores naranjas este empezará a tener más certeza de que es un elemento diferente. Este concepto cambia la forma en determinar una respuesta del sistema, en donde se basa en estadísticas para determinar inputs desconocidos.

⁴² AGUILAR, Anuor. ¿Qué es la tecnología cognitiva? [en línea]. Julio 2016. 24 marzo 2018. Disponible en <http://puntodemarketing.com/que-es-la-tecnologia-cognitiva/>

⁴³ Ibíd., p.1.

Estos sistemas son capaces de reconocer patrones, objetos y generar descripción de imágenes en lenguaje natural, un ejemplo de esto es el auto sin conductor el cual reconoce el entorno y es capaz entender la infinidad de entradas para así.

Volúmenes de información como libros de cursos universitarios pueden ser clasificados para extraer su información más fácilmente, por medio del análisis semántico de las palabras y clasificación de los resultados se puede lograr a dar información específica de manera más rápida.

3.2.2.1 WATSON

“Watson de IBM es la plataforma de inteligencia artificial más avanzada en el momento”⁴⁴, la cual recurre a un conjunto de tecnologías que utilizan el lenguaje natural y el aprendizaje automático como proceso similar al de los seres humanos, a diferencia de una base de datos que siempre da la misma respuesta Watson es capaz de aprender y ajustar su respuesta al entorno revelando información clave de grandes cantidades de datos estructurados y no estructurados, los datos estructurados son aquellos cuya información se encuentra ordenada y relacionada permitiendo su fácil comprensión por las herramientas de análisis, los datos no estructurados son todo lo contrario pero en ocasiones contienen más información que los datos estructurados.

Watson es capaz de ampliar sus conocimientos al interactuar con los humanos y añadir información en su base de datos, esta herramienta lo que busca es optimizar los procesos y hacer más corta la toma de decisiones, contestando las preguntas más complejas al extraer información de documentos extensos y de la experiencia de estas interacciones.

3.2.2.2 WATSON DISCOVERY

⁴⁴ IBM Watson. La tecnología cognitiva que abre una nueva era de la computación. [en línea]. Julio 2016. 24 marzo de 2018. Disponible en <http://puntodemarketing.com/que-es-la-tecnologia-cognitiva/>

Definido por IBM⁴⁵, Watson Discovery es uno de los catorce servicios de tecnología cognitiva de IBM, este se basa en la exploración de documentos en lenguaje natural para encontrar información, relacionarla y poder tomar decisiones con esta. Esta herramienta utiliza Análisis de datos combinado con Intuición Cognitiva (basada en NLP) para enriquecer data no estructurada.

Figura 9. características Watson Discovery.⁴⁶



El procesamiento se basa en.

- Subir, convertir, enriquecer y normalizar datos.
- Aplicar enriquecimiento adicional como conceptos, relaciones y sentimientos a través del Natural Language Understanding (NLU).
- Realizar búsqueda en lenguaje natural

El servicio está alojado en la nube y cuenta con un procesamiento de grandes cantidades de documentos.

⁴⁵ IBM. About Discovery. [en línea]. Febrero 2018. 27 marzo de 2018. Disponible en. <https://console.bluemix.net/docs/services/discovery/index.html#about>

⁴⁶ Ibíd., p.1.

3.2.2.3 PROCESAMIENTO DE LENGUAJE NATURAL

NLP es una rama de la inteligencia artificial que estudia las interacciones entre las máquinas y las personas por medio del habla común, según SAS⁴⁷ “aplicando algoritmos de entrenamiento de maquina tokenización y análisis sintáctico, tematización/ derivación, etiquetado de parte del habla, detección de lenguaje e identificación de relaciones semánticas”.

Actualmente procesar lenguaje natural implica un proceso de modelización matemática, los cuales hay dos grandes.

- **Modelos lógicos.** gramáticas que utilizan reglas y patrones estructurales creando diccionarios computacionales, normalmente buscando reflejar la estructura lógica del lenguaje.
- **Modelos Probabilísticos.** recoger colecciones de ejemplos para calcular frecuencias de unidades lingüísticas y la probabilidad de aparecer en un contexto, basados en aprendizaje automático.

Los componentes que se utilizan para entender la forma de las oraciones y frases son.

- **Análisis morfológico o léxico.** Consiste en el análisis interno de las palabras que forman oraciones para extraer lemas, rasgos flexivos, unidades léxica compuestas. Es esencial para la información básica. categoría sintáctica y significado léxico.

⁴⁷ SAS (s.f.). Natural Language Processing. [en línea]. 27 marzo 2018. Disponible en https://www.sas.com/es_co/insights/analytics/what-is-natural-language-processing-nlp.html

- **Análisis sintáctico.** Consiste en el análisis de la estructura de las oraciones de acuerdo con el modelo gramatical empleado (lógico o estadístico).
- **Análisis semántico.** Proporciona la interpretación de las oraciones, una vez eliminadas las ambigüedades morfosintácticas.
- **Análisis pragmático.** Incorpora el análisis del contexto de uso a la interpretación final. Aquí se incluye el tratamiento del lenguaje figurado (metáfora e ironía) como el conocimiento del mundo específico necesario para entender un texto especializado.

3.2.2.4 MODELOS

Un modelo se puede definir como una representación sencilla de la realidad el cual contiene sus características y propiedades permitiendo estudiar sus posibles comportamientos, para lo cual se hace necesario que estas características sean lo más simétricas o parecidas a las de la realidad que deseamos modelar.

Los modelos se construyen para conocer o predecir propiedades del objeto real, “la relación de correspondencia entre el objeto real y el modelo debe ser al menos parcialmente reversible y debe permitir la traducción de algunas propiedades del modelo a la realidad”⁴⁸, esto nos permite que un estado relativo al modelo pueda traducirse al modelo real y de esta forma que las repuestas derivadas del modelo. Sean aplicables a la realidad sin perder sentido.

La buena selección y descripción de la relación de los factores del objeto real se hace crucial para permitir un mayor conocimiento y predicción por parte del modelo construido, lo que da certeza a los resultados que se puedan generar con las diferentes interacciones con el modelo; cuando construimos un modelo construimos un sistema cuyos componentes derivados de la realidad se reducen a una cantidad manejable.

⁴⁸FELICISIMO, Manuel. Conceptos básicos, modelos y simulación [en línea] Disponible en Internet. http://www6.uniovi.es/~feli/CursoMDT/Tema_1.pdf.

Diferentes modelos pueden hacer representación al mismo fenómeno, objeto o sistema de la realidad, la diferencia se encuentra en la perspectiva con la que se creó cada uno, así como el detalle y descripción funcional de sus componentes

3.2.2.5 MODELO DE ENTRENAMIENTO DE MAQUINAS

Los modelos ⁴⁹consisten en algoritmos de entrenamiento de máquina que se dividen en tres categorías. supervisados, no supervisados y aprendizaje reforzado. Estos modelos necesitan de un set de entrenamiento para tener criterio en clasificar las entradas

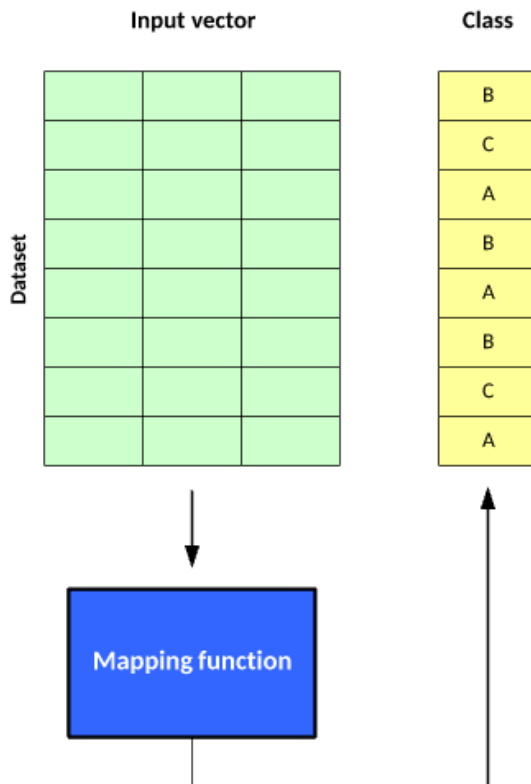
Aprendizaje supervisado. Son aquellos modelos que contienen una retroalimentación indicando si la respuesta está correcta o no, esta retroalimentación es revisada e ingresada en el set de datos de entrenamiento para alterar la respuesta del modelo en las próximas interacciones.

El set de datos se divide en dos fases. el set de datos y el set de entrenamiento, con el fin de sacar un nivel medible del rendimiento del modelo siendo el set de entrenamiento los datos que el modelo desconoce y debe predecir. Los algoritmos más comunes son support vector machines y el naïve Bayes

Aprendizaje no supervisado. Estos modelos no incluyen en set de datos una respuesta deseada, en donde clasifica los datos para que cada una de las clasificaciones tenga parte del set con características comunes

⁴⁹ IBM. Models for machine learning. [en línea]. Diciembre 2017. 27 marzo 2018. Disponible en developerWorks®. <https://www.ibm.com/developerworks/library/cc-models-machine-learning/index.html>

Figura 10. Aprendizaje no supervisado⁵⁰

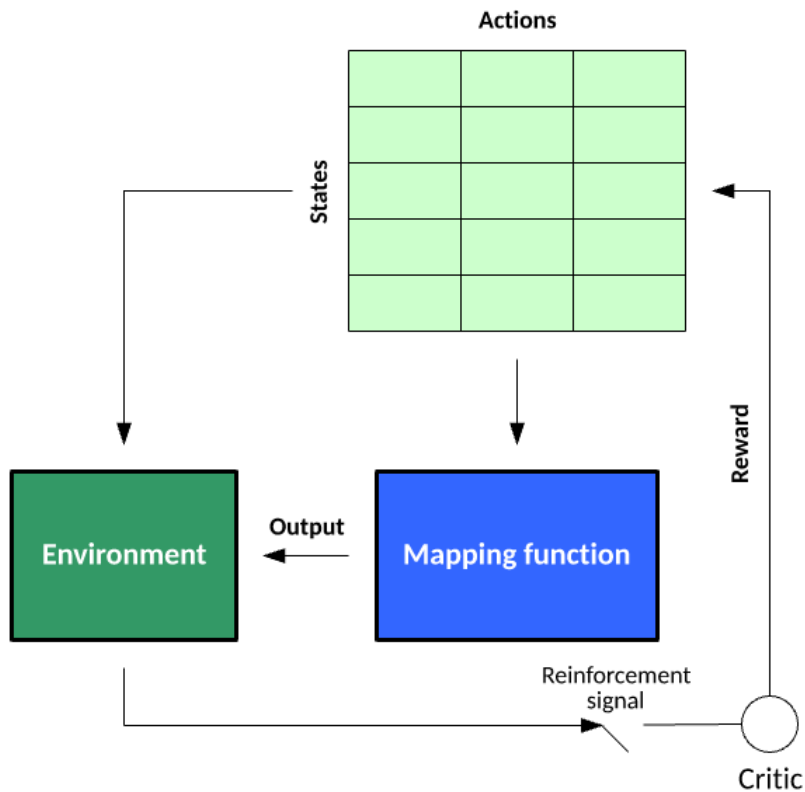


Las clases son generadas por la función y asigna a cada dato del set las características de la clase, por medio de un vector de características para clasificarlas, los algoritmos más comunes son el k-means clustering o el adaptive resonance theory.

Aprendizaje reforzado. Mantiene el esquema de las clasificaciones como los aprendizajes no supervisados, con la característica de proporcionar un cambio en el modelo cuando ocurre un estado específico, sin estar ligado a todas las acciones, solo aquellas que se les parametriza cambiar modificaran el modelo al llegar a tal estado.

⁵⁰ Ibíd., p.3.

Figura 11. Aprendizaje reforzado⁵¹



Análisis de datos. El análisis de datos consiste en un proceso de mejorar la calidad del set de datos, estandarizar los documentos en un formato en que el proceso lo entienda, para esto se utilizan.

3.3. MARCO DE ANTECEDENTES

3.3.1 HERRAMIENTAS DE CREACION

⁵¹ Ibíd., p.4.

Actualmente en el mercado existen diferentes herramientas para crear chatbots de forma dinámica y sin tener conocimientos en programación, se revisarán dos herramientas que pueden ser utilizadas para la creación.

3.3.1.1 DIALOGFLOW

Herramienta creada por Google para la interacción entre humanos y computadoras en lenguaje natural por medio de NLP. Cuenta con una interfaz de usuario amigable e intuitiva en la cual se proveen ejemplos de las formas en que los usuarios interactúan sobre un tema, se analiza y se entiende la finalidad de la pregunta.

Figura 12. Intenciones en DialogFlow⁵²

The screenshot displays the DialogFlow console for an intent named 'weather'. At the top, there is a 'SAVE' button and a menu icon. Below this, the 'Contexts' section is visible with a dropdown arrow. The 'User says' section includes a search bar labeled 'Search in user says'. A list of user expressions is shown, with the first one, 'Weather forecast in San Francisco tomorrow', selected and expanded. This expansion reveals a table of extracted entities:

PARAMETER NAME	ENTITY	RESOLVED VALUE
geo-city	@sys.geo-city	San Francisco
date	@sys.date	tomorrow

Below the table, other user expressions are listed: 'Weather for tomorrow', 'what is the weather today', and 'weather forecast'.

⁵² Google LLC Dialogflow Docs. [En línea] 21 02 2018. [Citado el. 17 03 2018.]<https://dialogflow.com/docs>.

⁵³la herramienta se basa en crear intenciones, un grupo de ejemplos que tienen una respuesta final en común y entidades, palabras claves necesarias para desambiguar una pregunta. La herramienta puede desactivar los algoritmos de Machine Learning para no usar el procesamiento de lenguaje natural.

Figura 13. Diálogos en DialogFlow⁵⁴

search_for_booking SAVE

Contexts ?

Events ?

Training phrases ? Search in user says

Add user expression

Book a hotel in Mountain View, check in on 11/03/2018 and check out on 11/08/2018

Action & parameters ?

book.hotel

REQUIRED ?	PARAMETER NAME ?	ENTITY ?	VALUE	IS LIST ?
<input type="checkbox"/>	destination	@sys.address	\$destination	<input type="checkbox"/>
<input type="checkbox"/>	check-in-date	@sys.date	\$check-in-date	<input type="checkbox"/>
<input type="checkbox"/>	check-out-date	@sys.date	\$check-out-date	<input type="checkbox"/>
<input type="checkbox"/>	Enter name	Enter entity	Enter value	<input type="checkbox"/>

+ New parameter

Para dar respuesta al usuario se crean diálogos, son flujos con condiciones lógicas que, dependiendo de la pregunta del usuario, las intenciones y entidades creadas se da una respuesta específica.

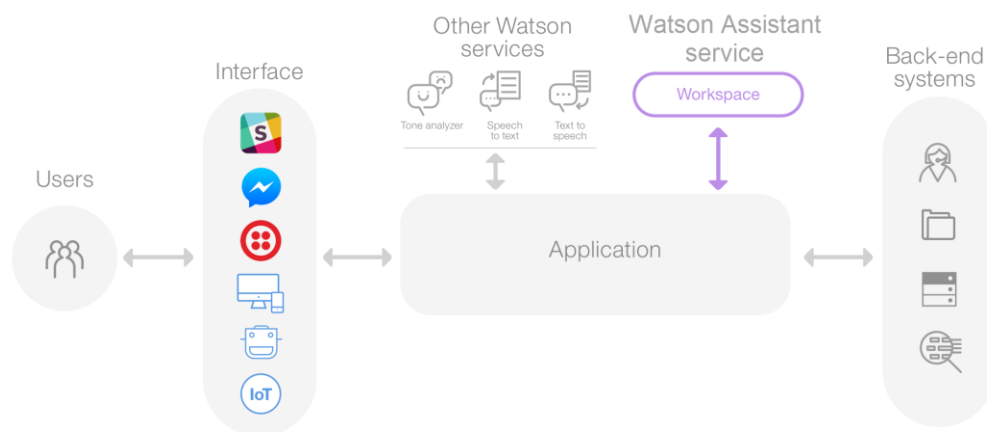
⁵³ Ibid., P.1.

⁵⁴ Ibid., P.1.

3.3.1.2 WATSON ASSISTANT

⁵⁵Desarrollado como SaaS, es una de los 15 servicios de Watson IBM para construir soluciones de entendimiento de lenguaje natural en texto usando aprendizaje de máquina.

Figura 14. Ej. Estructura de Implementación de Watson Assistant⁵⁶



La herramienta se maneja por medio de Workspaces divididos en tres ramas. intenciones, entidades y diálogos. Las intenciones son acciones que el usuario requiere y se entrenan con formas de preguntar que tienen la misma finalidad, las entidades son palabras claves que se utilizan para dar una respuesta final más acorde específica, los diálogos con flujos condicionados que tomas los dos primeros puntos y condiciones lógicas para dar la respuesta más correcta.

⁵⁵ **IBM**. About Watson Assistant. [En línea]26 01 2018. [Citado el. 17 03 2018.]<https://console.bluemix.net/docs/services/conversation/index.html#about>.

⁵⁶Ibíd., P.1.

La herramienta en las aplicaciones es consumida, este retorna la respuesta más lógica y la aplicación decide que utilizar de la respuesta.

3.3.2 PRODUCTOS BASADOS EN CHATBOTS

Existen actualmente en el mercado productos del sector educativo, basados en brindar información de las instituciones educativas y los servicios que ofrecen.

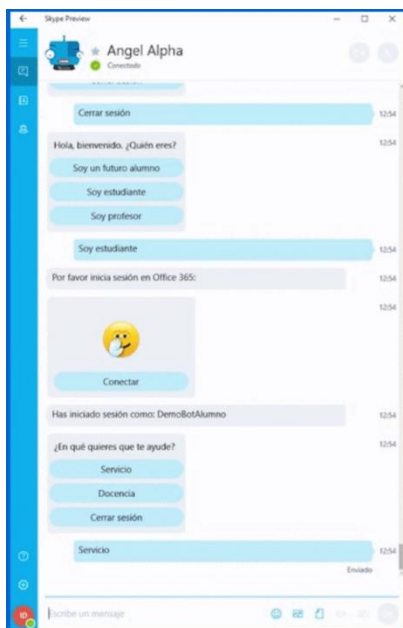
3.3.2.1 EDUARBOT

⁵⁷Encamina y Microsoft Bot Framework desarrollaron un chatbot asistente virtual para los alumnos, brindando información de los servicios del centro de educación con la posibilidad de conectar con los servicios de los centros educativos brindando información de Carreas, Administración, Bibliotecas, Deportes, etc.

Figura 15. Ejemplo EduarBot⁵⁸

⁵⁷ **ENCAMINA.** Revoluciona tu centro educativo con EduarBot. [En línea][Citado el. 17 03 2018.]<https://www.encamina.com/eduarbot/>.

⁵⁸ *Ibíd.*, P.1.



Con el soporte de Cognitive Services de Azure, es capaz de entender charlas de forma informal e interpretar lo que necesita el alumno en diferentes canales como Skype o Messenger.

3.3.2.2 WANDA

⁵⁹Unit4, desarrolladora de software empresarial junto con Microsoft Azure desarrollo un chatbot para los estudiantes de educación superior utilizando componentes de aprendizaje automático, análisis de flujo y procesamiento de eventos de office 365, brinda información de gestión de cursos, asignación de citas e información de la institución con integraciones multicanal.

Figura 16. Ejemplo Wanda⁶⁰

⁵⁹ **García, Óscar.** Unit4 presenta su primer chatbot para el estudiante de Educación Superior. [En línea]25 10 2016. [Citado el. 17 03 2017.]<https://www.unit4.com/es/acerca-de/noticias/2016/10/unit4-presenta-su-primer-chatbot-para-el-estudiante-de-educacion-superior>.

⁶⁰ Ibíd. P.,1.



Con una arquitectura completamente PaaS esta herramienta consume los servicios de Microsoft Azure y Unit4 Student Management para ser integrada en diferentes canales como Facebook Messenger, Skype o WeChat.

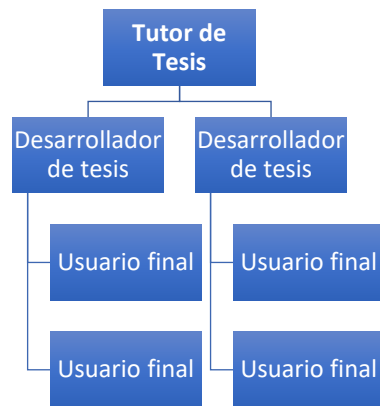
4. PLAN DE DESARROLLO

4.1. IDENTIFICACIÓN DE INVOLUCRADOS

Tabla 1. Involucrados del proyecto (Autores).

Nombre	Cargo	Rol en el proyecto	Correo
Gilberto Pedraza García	Profesor de la escuela TIC	Tutor de Tesis	
Gustavo Iván Arce Millán	Estudiante de la escuela TIC	Desarrollador de tesis	gustavo-arce@upc.eu.co
Omar David Sánchez Ortiz	Estudiante de la escuela TIC	Desarrollador de tesis	omar-sanchez@upc.edu.co
Muestra	Estudiantes	Usuario final.	N/A

Figura 16. Jerarquía de involucrados (Autores).



4.2. ALCANCE Y DEFINICIÓN DE TRABAJO

Tabla 2. Definición de actividades (Autores).

Actividad	Entregables	Fecha de Entrega x Sprint
<ul style="list-style-type: none"> Definir objetivo general, Definir objetivos específicos Definir limitar el proyecto Argumentar proyecto Hipótesis Variables 	<ul style="list-style-type: none"> Objetivos del proyecto Límites del proyecto Justificación del proyecto Hipótesis del proyecto con las variables dependientes y dependientes que la afectan 	Sprint 1 Semana (1-2)
<ul style="list-style-type: none"> Desarrollo Marco teórico Desarrollo Marco conceptual Marco de Antecedentes Cronograma de Actividades 	<ul style="list-style-type: none"> Marco referencial Marco de antecedentes Cronograma de actividades 	Sprint 2 Semana (3-4)
<ul style="list-style-type: none"> Planteamiento de la solución. Levantamiento de requerimientos. 	<ul style="list-style-type: none"> Pool de soluciones Selección de solución más acertada Historias de Usuario Requerimientos funcionales Requerimientos no funcionales 	Sprint 3 Semana (5-6)
<ul style="list-style-type: none"> Diseño de prototipo (Arquitectura de software, diagramas UML). 	<ul style="list-style-type: none"> Diagrama de secuencia Diagrama de clases Diagrama de casos de uso Diagrama de secuencia Diagrama de base de datos 	Sprint 4 Semana (7-8)
<ul style="list-style-type: none"> Desarrollo de la solución. Backend. 	<ul style="list-style-type: none"> Aprovisionamiento de Base de datos. Modelo de datos. 	Sprint 5 Semana (9-10)
<ul style="list-style-type: none"> Front-end Aprovisionamiento de Watson Discovery. Data crawler. 	<ul style="list-style-type: none"> Interfaz inicial Data crawler. 	Sprint 6 Semana (11-12)
<ul style="list-style-type: none"> Parametrización de la herramienta Pruebas unitarias 	<ul style="list-style-type: none"> Integración de la solución. 	Sprint 7 Semana

•	• Capítulo de pruebas (13-14)
• Pruebas de aceptación de usuario	• Capítulo conclusiones y recomendaciones.
• Análisis de resultado	• Software

4.3. PLAN DE COMUNICACIÓN

Tabla 3. Plan de comunicación (Autores).

	Miembros del equipo	Tutor de tesis	Usuarios
Plan de trabajo	x		
Desarrollo de soluciones	x		
Pruebas de soluciones	x	x	
Despliegue de soluciones	x	x	x
Resumen ejecutivo	x	x	
Riesgos / supuestos	x	x	
Ceremonias			
Daily meeting	X		
Planning meeting	X		
Review meeting	x		
Demo meeting – x reléase	x	x	

Se busca informar al tutor de tesis cuando los aplicativos estén en a un nivel de presentación.

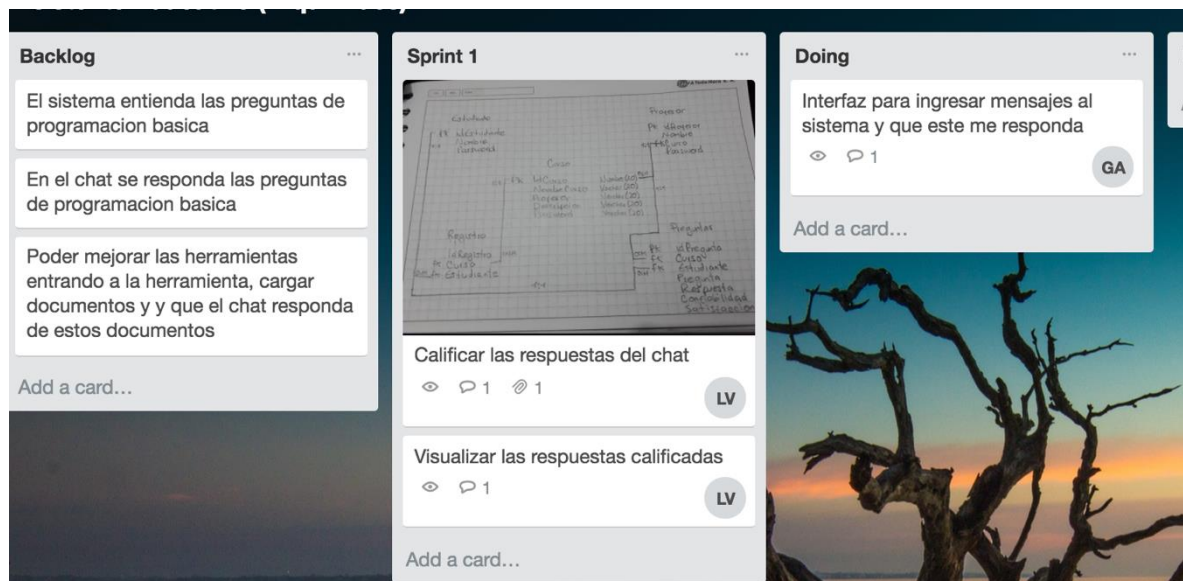
4.4. CRONOGRAMA

Tabla 4. Cronograma (Autores).

Actividad	Entregables	Fecha de Entrega x Sprint
<ul style="list-style-type: none"> Definir objetivo general, Definir objetivos específicos Definir limitar el proyecto Argumentar proyecto Hipótesis Variables 	<ul style="list-style-type: none"> Objetivos del proyecto Límites del proyecto Justificación del proyecto Hipótesis del proyecto con las variables dependientes y dependientes que la afectan 	Sprint 1 Semana (1-2)
<ul style="list-style-type: none"> Desarrollo Marco teórico Desarrollo Marco conceptual Marco de Antecedentes Cronograma de Actividades 	<ul style="list-style-type: none"> Marco referencial Marco de antecedentes Cronograma de actividades 	Sprint 2 Semana (3-4)
<ul style="list-style-type: none"> Planteamiento de la solución. Levantamiento de requerimientos. 	<ul style="list-style-type: none"> Pool de soluciones Selección de solución más acertada Historias de Usuario 	Sprint 3 Semana (5-6)
<ul style="list-style-type: none"> Diseño de prototipo (Arquitectura de software, diagramas UML). 	<ul style="list-style-type: none"> Diagrama de secuencia para tener una generalidad del comportamiento de la aplicación Diagrama de clases que ayuden la detección de funcionalidades Diagrama de casos de uso Diagrama de secuencia Diagrama de base de datos para detectar los datos a capturar 	Sprint 4 Semana (7-8)
<ul style="list-style-type: none"> Desarrollo de la solución. 	<ul style="list-style-type: none"> Aprovisionamiento 	Sprint 5

<ul style="list-style-type: none"> • Backend. 	<ul style="list-style-type: none"> • de Base de datos. • Modelo de datos. 	Semana (9-10)
<ul style="list-style-type: none"> • Front-end • Aprovisionamiento de Watson Discovery. • Data crawler. 	<ul style="list-style-type: none"> • Interfaz inicial • Data crawler. 	Sprint 6 Semana (11-12)
<ul style="list-style-type: none"> • Parametrización de la herramienta • Pruebas unitarias • 	<ul style="list-style-type: none"> • Integración de la solución. • Capítulo de pruebas 	Sprint 7 Semana (13-14)
<ul style="list-style-type: none"> • Pruebas de aceptación de usuario • Análisis de resultado 	<ul style="list-style-type: none"> • Capitulo conclusiones y recomendaciones. • Software 	

Figura 17. Trello (Autores).



5. DISEÑO DE PROTOTIPO

Al plantear la solución del problema, se empezó por aterrizar la propuesta en tareas que al ejecutarse se integran para conformar la solución, definiéndolas en funcionalidades.

Como se menciona en el capítulo de la metodología, se levantaron “historias de usuario” que ayudaron a dividir el total del producto en entregas parciales de unas funcionalidades específicas.

HISTORIA DE USUARIO			
NOMBRE		Interacción estudiante con chat.	
ID	HU-001	ACTOR	Estudiante
PRIORIDAD	Media	SPRINT	4
RIESGO DESARROLLO	EN Bajo	DESARROLLADOR	GUSTAVO IVAN
DESCRIPCION HISTORIA DE USUARIO			
<p>Como estudiante requiero interactuar por medio de texto con el chat y que este me responda temas alusivos al programa fundamentos de programación.</p>			
ENTRADAS		SALIDAS	
<ul style="list-style-type: none"> Pregunta en forma de texto por parte del usuario 		<ul style="list-style-type: none"> Respuesta del chat con bases en fundamentos de programación 	
PRECONDICIONES		POSTCONDICIONES	
<ul style="list-style-type: none"> Acceso del usuario a la página que contenga el chat. 		<ul style="list-style-type: none"> 	

DEPENDENCIA		EXTENSION	
•		• HU-002	
CRITERIOS DE ACEPTACIÓN			
CRITERIO DE ACEPTACION	CONTEXTO		
1	Responder con base a la pregunta ingresada.		
SEGUIMIENTO AL CASO DE USO			
Fecha	Descripción	Gestor	

HISTORIA DE USUARIO			
NOMBRE		Validación respuestas del chat por estudiante.	
ID	HU-002	ACTOR	Estudiante
PRIORIDAD	Media	SPRINT	4
RIESGO DESARROLLO	EN Bajo	DESARROLLADOR	GUSTAVO IVAN
DESCRIPCION HISTORIA DE USUARIO			
<p>Como estudiante necesito calificar como correcta o incorrecta la respuesta que el chat me dé para enviar una realimentación.</p>			

ENTRADAS		SALIDAS	
Respuesta de chat.		Confirmación de la calificación de la respuesta dada por parte del estudiante,	
PRECONDICIONES		POSTCONDICIONES	
HU-001		•	
DEPENDENCIA		EXTENSION	
HU-003		•	
CRITERIOS DE ACEPTACIÓN			
CRITERIO DE ACEPTACION		CONTEXTO	
1		Respuesta relacionada o alusiva a palabras claves en la pregunta del estudiante.	
SEGUIMIENTO AL CASO DE USO			
Fecha		Descripción	Gestor

HISTORIA DE USUARIO			
NOMBRE		Carga de Archivos en el modelo.	
ID	HU-003	ACTOR	Profesor
PRIORIDAD	Media	SPRINT	4
RIESGO EN DESARROLLO	Bajo	DESARROLLADOR	GUSTAVO IVAN
DESCRIPCION HISTORIA DE USUARIO			
<p>Como profesor requiero poder subir en el sistema los documentos que considere necesarios para que el sistema responda en base a estos.</p>			

ENTRADAS		SALIDAS
Documentación		Confirmación de carga de documentación.
PRECONDICIONES		POSTCONDICIONES
		•
DEPENDENCIA		EXTENSION
		•
CRITERIOS DE ACEPTACIÓN		
CRITERIO DE ACEPTACION	CONTEXTO	
1	se deben poder cargar documentos en formato pdf.	
2	Se deben visualizar respuestas basadas en estos nuevos documentos.	
SEGUIMIENTO AL CASO DE USO		
Fecha	Descripción	Gestor

HISTORIA DE USUARIO			
NOMBRE		Interacción profesor con Información de interacción	
ID	HU-004	ACTOR	Profesor
PRIORIDAD	Media	SPRINT	4
RIESGO DESARROLLO	EN Bajo	DESARROLLADOR	GUSTAVO IVAN
DESCRIPCION HISTORIA DE USUARIO			
<p>Como profesor requiero visualizar información descriptiva de las calificaciones dadas por los estudiantes a las respuestas del chat para validar el correcto funcionamiento del sistema y la documentación cargada.</p>			

ENTRADAS		SALIDAS
<ul style="list-style-type: none">Ingreso en el sistema como profesor.Petición de reporte de calificaciones.		<ul style="list-style-type: none">Calificación de las respuestasRespuestas dadas por el chatGrafica con información de las respuestas
PRECONDICIONES		POSTCONDICIONES
HU-001		<ul style="list-style-type: none">
DEPENDENCIA		EXTENSION
HU-001		<ul style="list-style-type: none">
CRITERIOS DE ACEPTACIÓN		
CRITERIO DE ACEPTACION	CONTEXTO	
1	los datos deben ser almacenados al cien por ciento de las interacciones	
2	los datos deben poderse consultar	
SEGUIMIENTO AL CASO DE USO		
Fecha	Descripción	Gestor

HISTORIA DE USUARIO			
NOMBRE		Interacción usuario con modelo	
ID	HU-005	ACTOR	Profesor
PRIORIDAD	Media	SPRINT	4
RIESGO EN DESARROLLO	Bajo	DESARROLLADOR	GUSTAVO IVAN

DESCRIPCION HISTORIA DE USUARIO		
<p>Como profesor administrador requiero poder renquear las respuestas que da el chat para que este de la respuesta más acertada.</p>		
ENTRADAS		SALIDAS
Pregunta hecha por el estudiante		Prioridad de respuestas
PRECONDICIONES		POSTCONDICIONES
HU-001, HU-004		<ul style="list-style-type: none">
DEPENDENCIA		EXTENSION
HU-001, HU-004		<ul style="list-style-type: none">
CRITERIOS DE ACEPTACIÓN		
CRITERIO DE ACEPTACION	CONTEXTO	
SEGUIMIENTO AL CASO DE USO		
Fecha	Descripción	Gestor

5.1. ARQUITECTURA DE SOFTWARE

Una vez concretadas todas las historias de usuario se tuvo un panorama más claro de lo que se quería llegar a desarrollar para involucrar todas las historias y saber que componentes tenían que implementarse, se crearon diagramas que mostraran las diferentes facetas del prototipo.

Figura 18. Vista de contexto (Autores)

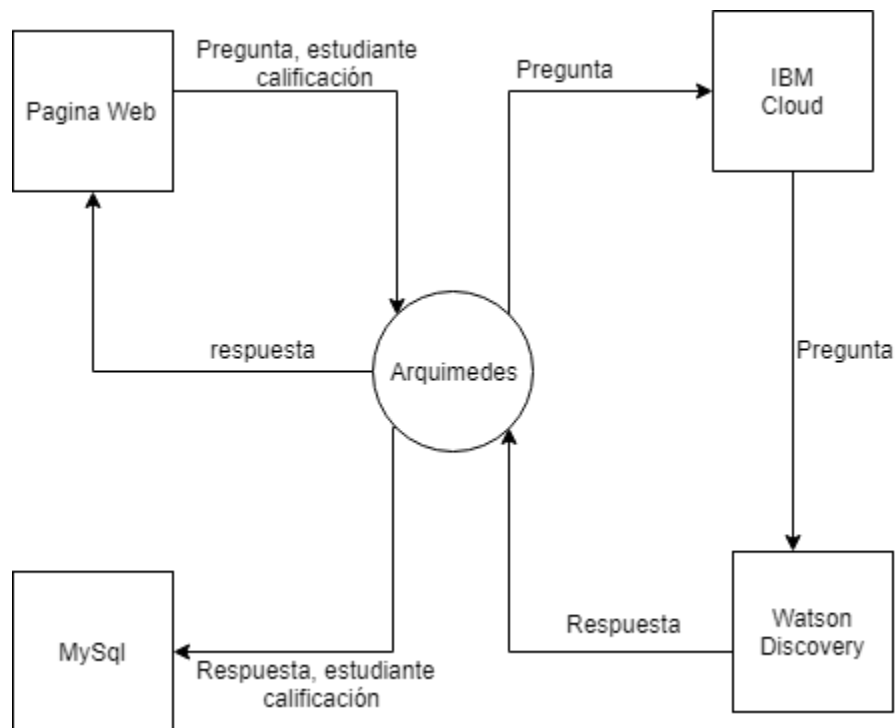
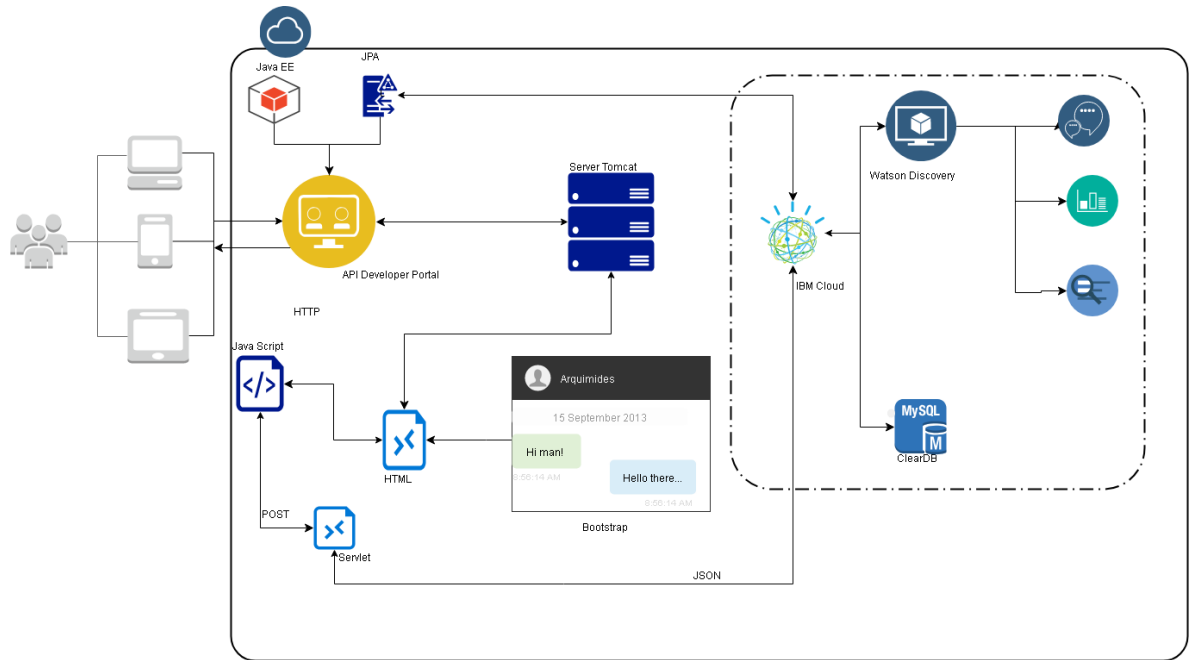


Figura 19. Arquitectura de Software (Autores))



El usuario accede a la aplicación que embeberá el chat a través de cualquier dispositivo. Esta aplicación contendrá el chat que está alojado en el servidor Tomcat y las preguntas que el usuario le haga las responderá a través del servicio Watson Discovery que se encuentra en la nube de IBM, sin embargo, este servicio se encuentra personalizado por los desarrolladores a través de diferentes técnicas (como análisis de datos y búsquedas enriquecidas) permitiendo no solo entregar una mejor experiencia y calidad de respuestas.

5.2. ESTRUCTURA FISICA DE LA SOLUCION

Al tener los bloques de componentes que hacen parte de la arquitectura se empezó a detallar el detalle de las conexiones entre cada componente, algunos están externos a la aplicación por la disponibilidad de los servicios que ofrecen.

Figura 20. Diagrama de Componentes Solución del Problema (Autores).

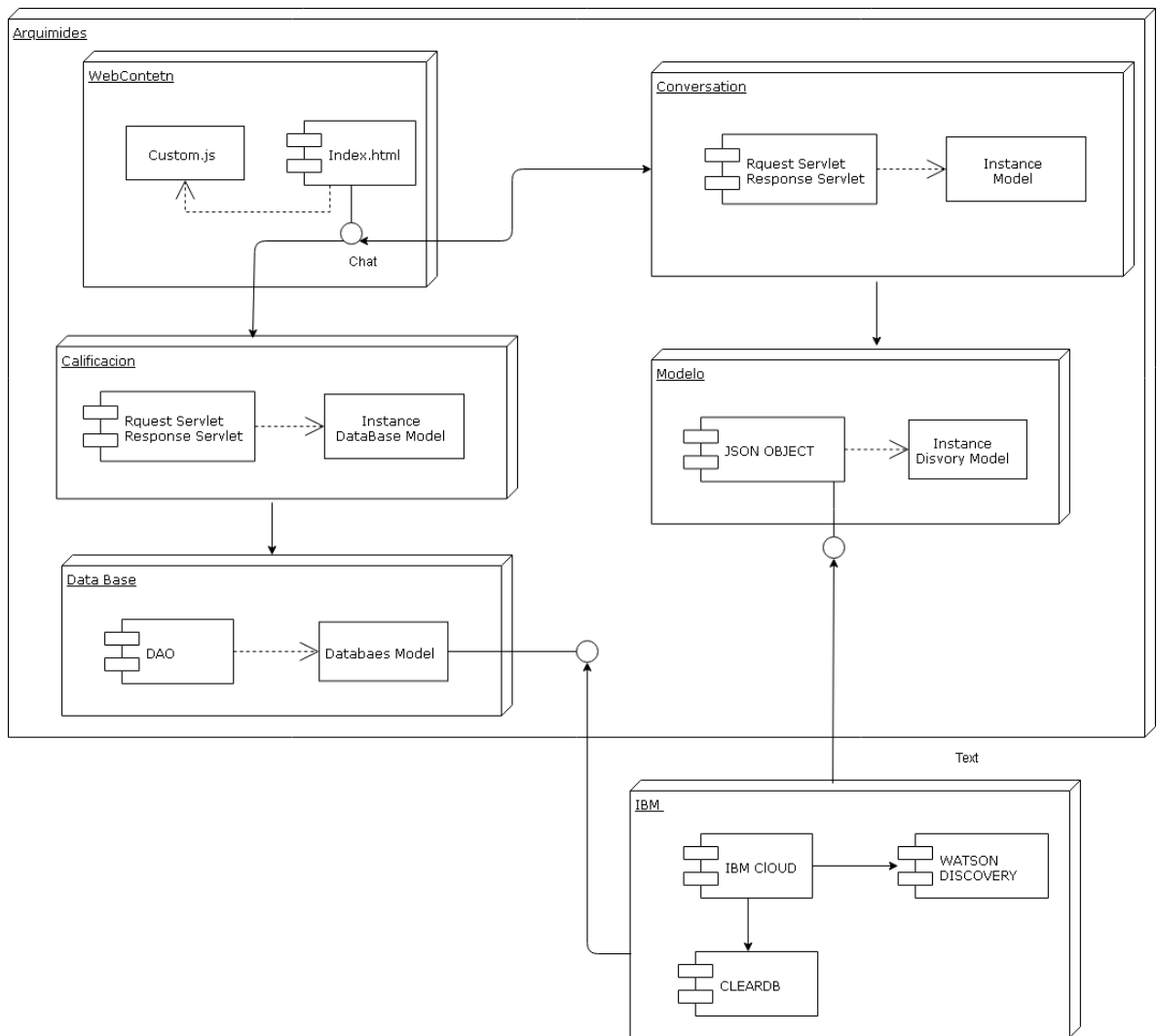
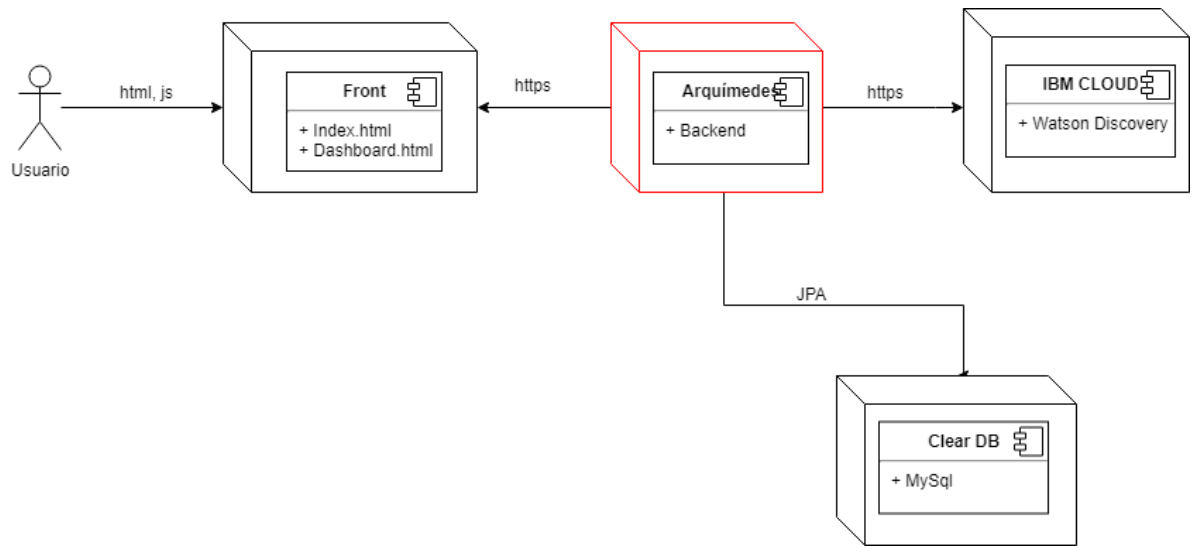


Figura 21. Despliegue (nivel de instancia-Autores)



5.2.1 Escenarios de Calidad

Escenario de Calidad #	1	Stakeholder.	Usuario
Atributo de Calidad	Disponibilidad		
Justificación	Mantener operativo el componente		
Fuente	Aplicacion		
Estímulo	Disponibilidad		
Artefacto	Sistema		
Ambiente	Normal		

Respuesta	Poder entrar a interactuar
LMedida de la Respuesta	Disponibilidad al 100

Escenario de Calidad #	2	Stakeholder.	Usuario
Atributo de Calidad	Seguridad		
Justificación	El sistema y sus componentes solo puede ser accedido en la página del curso		
Fuente	Aplicación		
Estímulo	Seguridad		
Artefacto	Sistema		
Ambiente	Producción		

Respuesta	Restringir el acceso a ambiente al que aplica
Medida de la Respuesta	la url de los componentes deben estar ocultos en el Backend y/o ocultados en el front-end

Escenario de Calidad #	3	Stakeholder.	Usuario
Atributo de Calidad	Concurrencia		
Justificación	poder responder simultaneas peticiones de diferentes usuarios		
Fuente	Aplicación		
Estímulo	Concurrencia		
Artefacto	Sistema		
Ambiente	Producción		

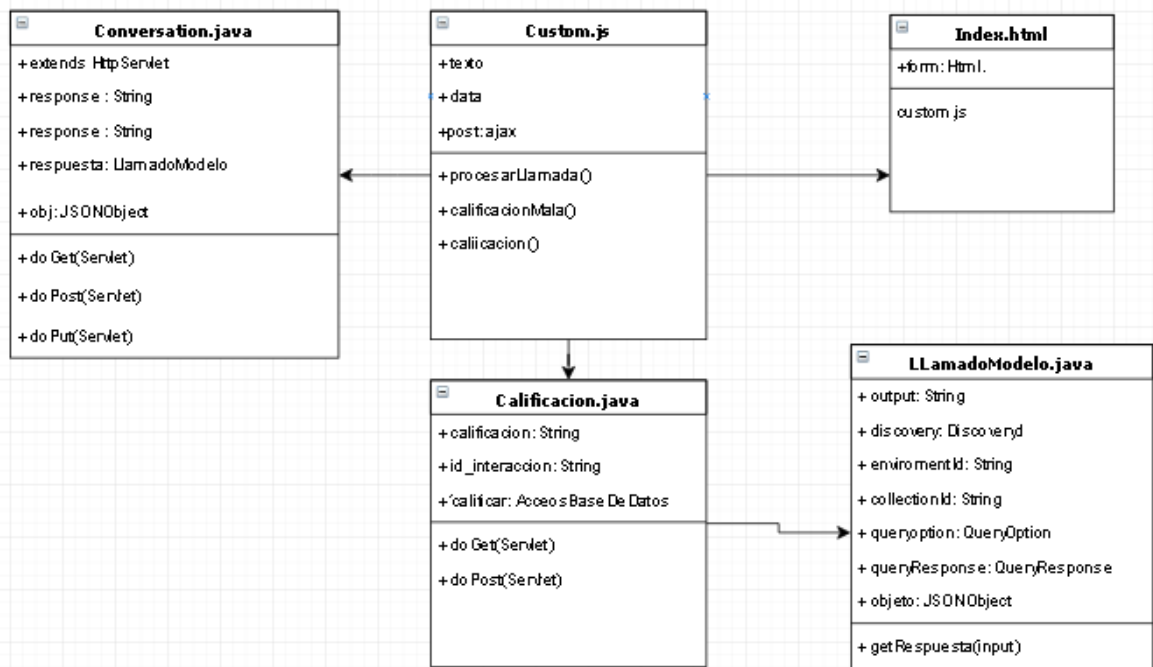
Respuesta	Pruebas de carga del sistema
Medida de la Respuesta	Establecer límite de concurrencias y medidas de abarcarla limitación

Escenario de Calidad #	4	Stakeholder.	Usuario
Atributo de Calidad	Tiempo de respuesta		
Justificación	la aplicación debe responder en un máximo de 5 segundos al usuario		
Fuente	Aplicación		
Estímulo	Tiempo de respuesta		
Artefacto	Sistema		
Ambiente	Producción		

Respuesta	Establecer tiempos de respuesta
Medida de la Respuesta	Optimización de los componentes

5.3. DIAGRAMA DE CLASES

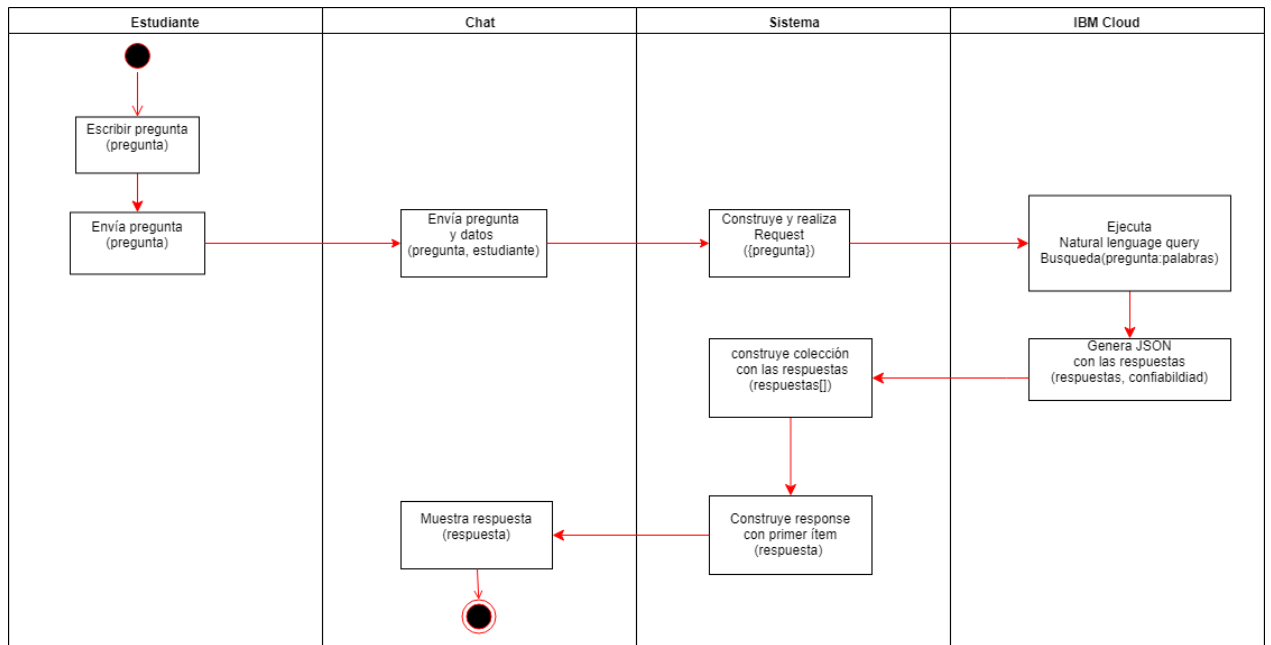
Figura 22. Diagrama de Clases (Autores).



5.4. INTERACCIONES SISTEMA

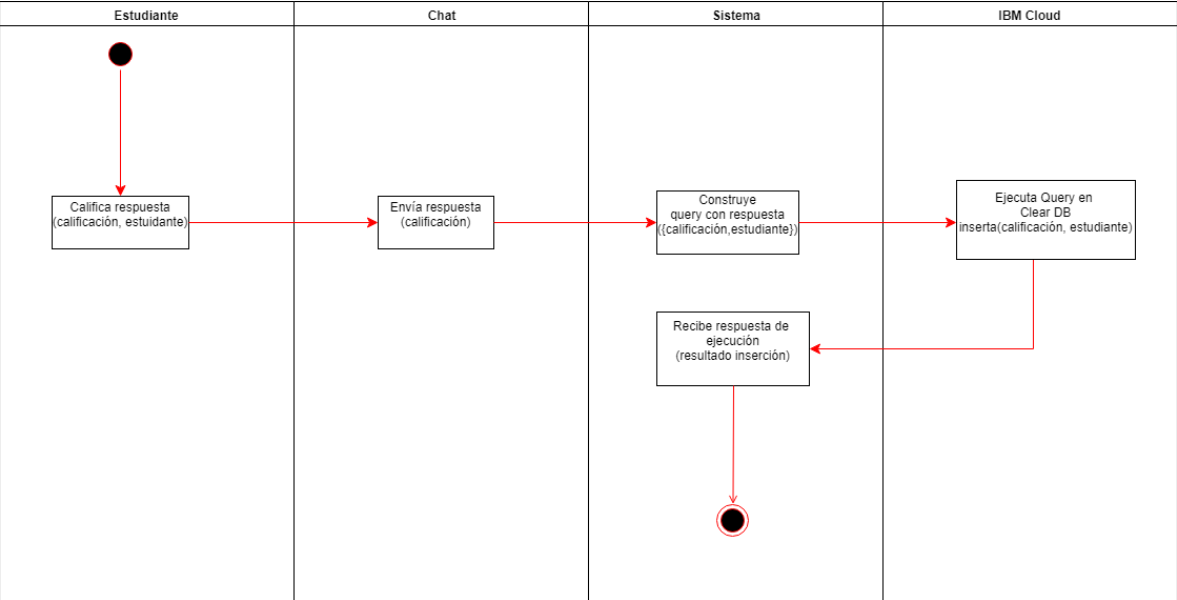
5.4.1 INTERACCION ESTUDIANTE-AULA VIRTUAL

Figura 23. Interacción estudiante Aula virtual (Autores).



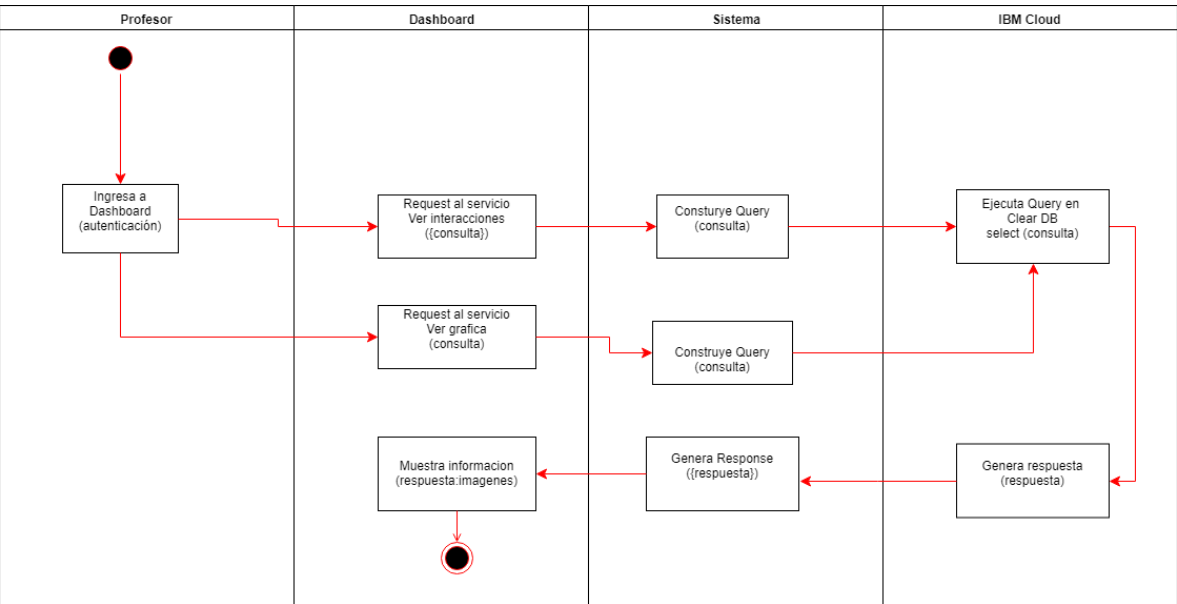
5.4.2 CALIFICACIÓN RESPUESTAS

Figura 24. Diagrama calificación respuestas (Autores).



5.4.3 VISUALIZACIÓN PREGUNTAS Y RESPUESTAS

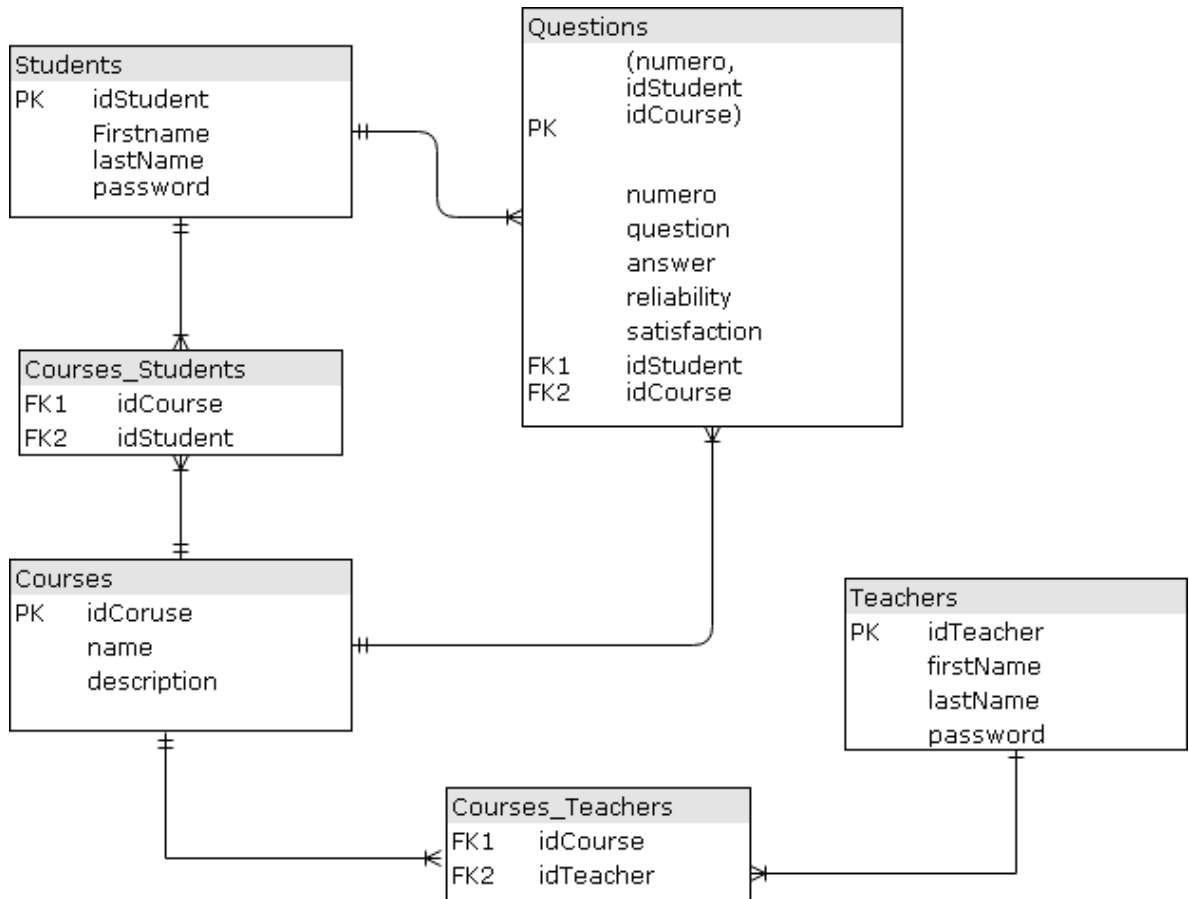
Figura 25. Diagrama visualización preguntas y respuestas (Autores).



5.5. BASE DE DATOS

5.5.1 DIAGRAMA DE BASE DE DATOS

Figura 26. Diagrama base de datos. (Autores).



5.5.2 SCRIPTS

5.5.2.1 CREACION BASE DE DATOS

- Se crea la entidad Estudiantes.
`CREATE TABLE students (idStudent INT auto_increment PRIMARY KEY, firstName VARCHAR(50), lastName VARCHAR(50), email VARCHAR(50), password VARCHAR(50));`
- Se crea la entidad Profesores.
`CREATE TABLE teachers (idTeacher INT auto_increment PRIMARY KEY, firstName VARCHAR(50), lastName VARCHAR(50), email VARCHAR(50), password VARCHAR(50));`
- Se crea la entidad Cursos.
`CREATE TABLE courses (idCourse INT auto_increment PRIMARY KEY, name VARCHAR(50), description VARCHAR(50));`
- Se crea la relación entre estudiantes y cursos.
`CREATE TABLE courses_students (idCourse int, idStudent int);`
- Se crea la llave foránea idCurso en la tabla courses_students
`ALTER TABLE courses_students ADD FOREIGN KEY(idCourse) REFERENCES courses(idCourse);`
- Se crea la llave foránea idStudent en la tabla courses_students
`ALTER TABLE courses_students ADD FOREIGN KEY(idStudent) REFERENCES students(idStudent);`
- Se crea la relación entre profesores y cursos.
`CREATE TABLE courses_teachers (idCourse INT, idTeacher INT);`
- Se crea la llave foránea idCourse en la tabla courses_teachers
`ALTER TABLE courses_teachers ADD FOREIGN KEY(idCourse) REFERENCES courses(idCourse);`

- Se crea la llave foránea idTeacher en la tabla courses_teachers
`ALTER TABLE courses_teachers ADD FOREIGN KEY(idTeacher) REFERENCES teachers(idTeacher);`
- Se crea la entidad Preguntas
`CREATE TABLE questions (numero INT auto_increment, idCourse INT, idStudent INT, question VARCHAR(400), answer VARCHAR(400), reliability VARCHAR(400), satisfaction INT(1), PRIMARY KEY (numero,idCourse,idStudent));`
- Se crea la llave foránea idStudent en la tabla questions
`ALTER TABLE questions ADD FOREIGN KEY(idStudent) REFERENCES students(idStudent);`
- Se crea la llave foránea idCourse en la tabla questions
`ALTER TABLE questions ADD FOREIGN KEY(idCourse) REFERENCES courses(idCourse);`

6. IMPLEMENTACION

6.1. ESPECIFICACIONES TÉCNICAS

La aplicación Arquímedes es una aplicación web JEE utilizando la versión 1.8 del JRE, implementa la clase Servlet para mapear los servicios web y JPA para el manejo de persistencia en la base de Datos, esta implementada para servidores Apache Tomcat 8 aunque es posible implementarse en otros servidores de aplicaciones con características similares.

6.2. FRONT-END

El Front-End de la aplicación se desarrolló en HTML y JavaScript implementando la biblioteca JQuery y Bootstrap en donde se hacen llamados HTTP para recibir información del Back-End por medio de peticiones POST.

La parte visual del contenido web está alojado en el archivo **Index.html**, este a su vez implementa el archivo **Custom.js** que se encarga de administrar la lógica de la página, crear los componentes dinámicos y capturar los datos del Back-End.

6.3. BACK-END

Los Servlets expuestos solo reciben peticiones en método POST con la información en formato JSON, estos se encargan de transformar la información recibida y enviarla a las clases pertinentes y responder las peticiones.

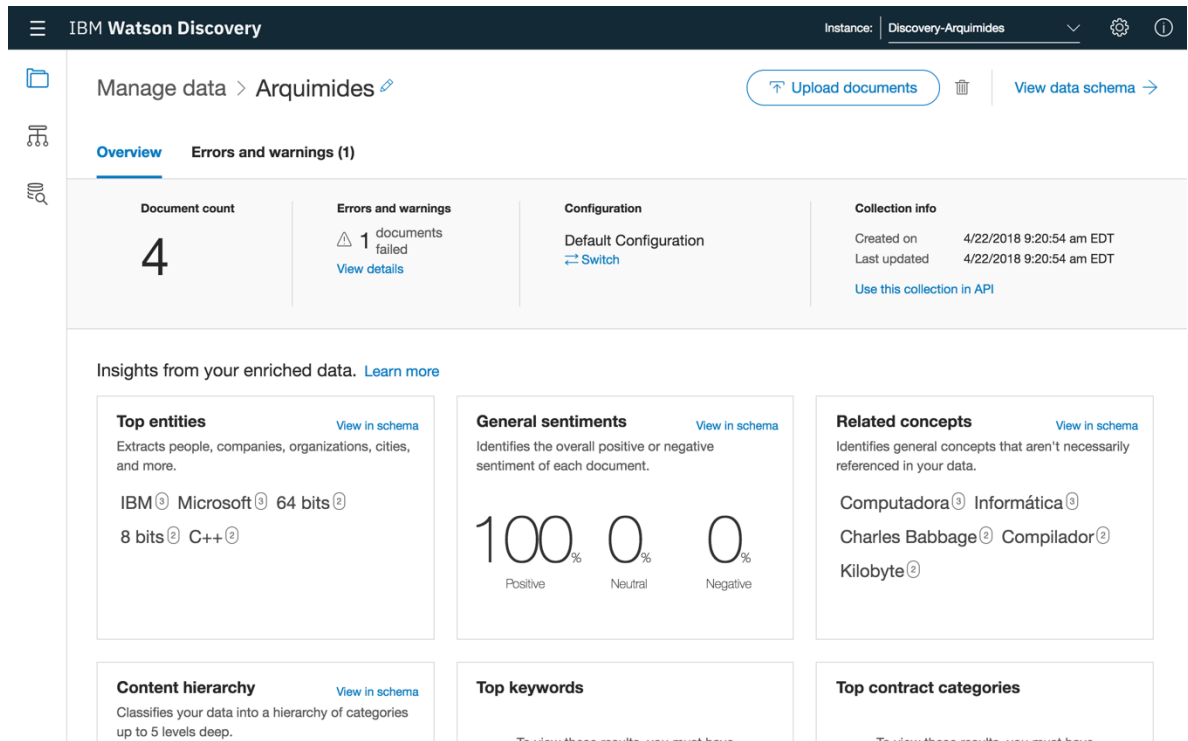
Los componentes de la aplicación hacen llamados a API's alojadas en la infraestructura IBM Cloud por medio del JDK de Watson IBM.

Utiliza JPA para mapear las entidades y persistir en una instancia de Base de datos relacional MYSQL llamada ClearDB, alojada en la infraestructura IBM Cloud.

6.4. WATSON DISCOVERY

Esta herramienta está alojada en la infraestructura IBM Cloud y expone una herramienta visual en línea para la carga de documentos, análisis, normalización de datos y refinamiento de búsquedas.

Figura 27. Interfaz Watson Discovery (Autores).



Esta herramienta utiliza algoritmos de Entrenamiento de máquina y técnicas de estandarización de datos, toda la herramienta consume servicios del API que pueden también ser consumidos por aplicaciones, la finalidad de la forma visual es que un usuario funcional pueda interactuar con los datos, refinar las búsquedas y mejorar las respuestas del servicio.

6.5. DIAGRAMAS FRONT-END

Figura 28. Diagrama Front-End Chat sin desplegar (Autores).

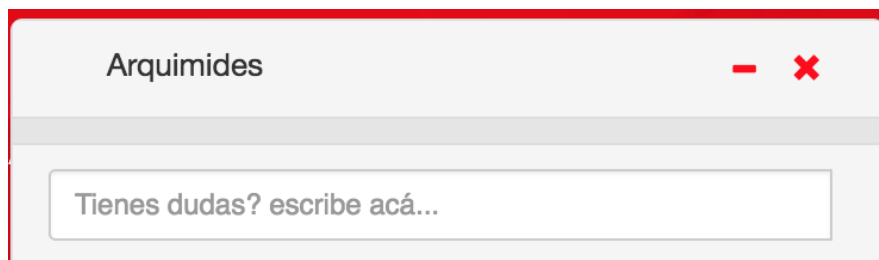


Figura 29. Diagrama Front-End Chat Desplegado (Autores).

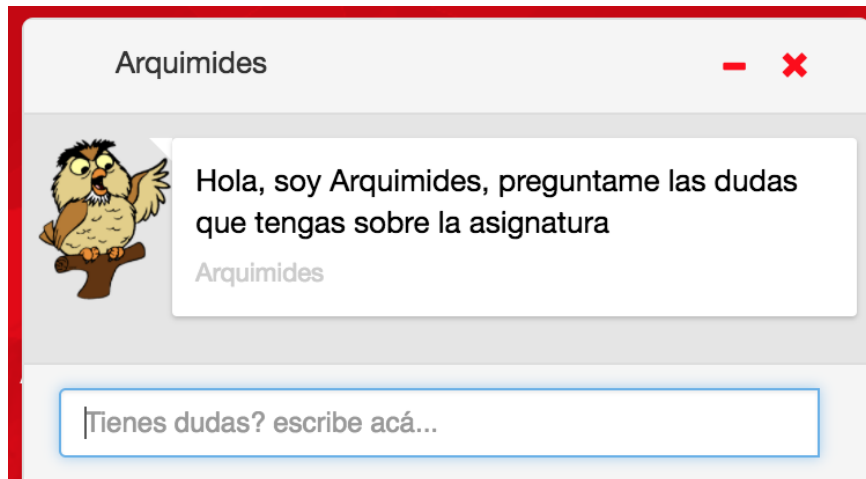
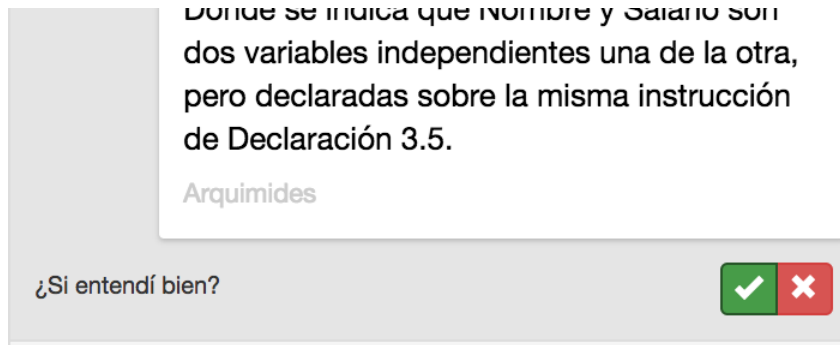


Figura 30 Diagrama Front-End Respuesta del Chat (Autores).



Figura 31. Diagrama Front-End Calificación (Autores).



6.6. DIAGRAMAS FRONT END

Figura 32. Diagrama Back-End Interacción (Autores).

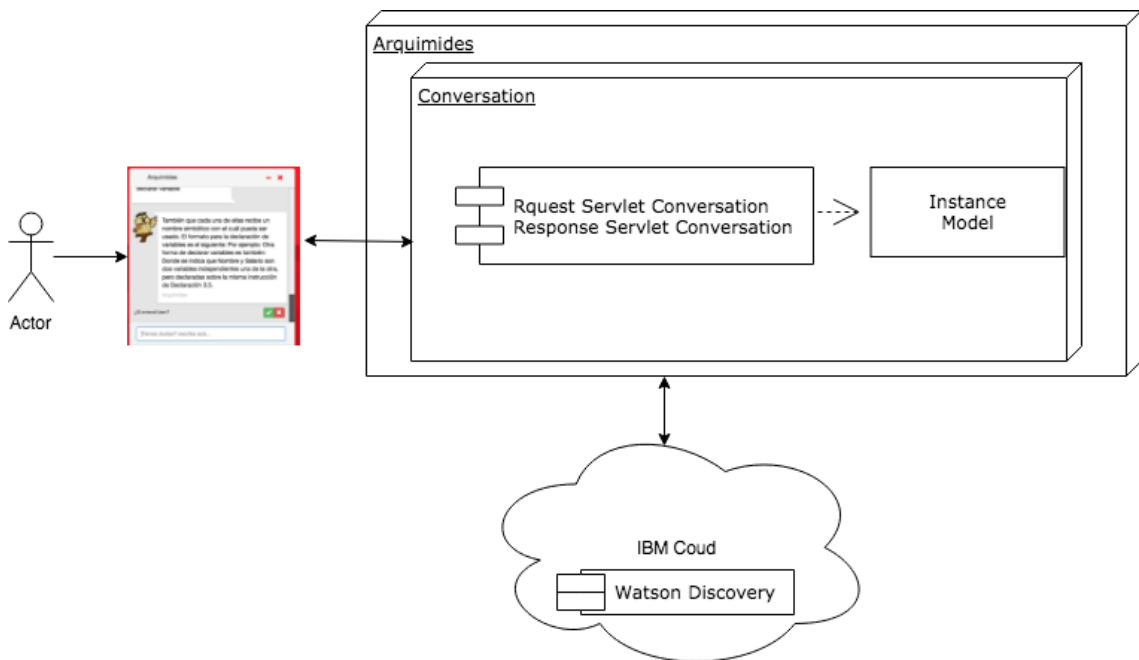


Figura 33. Diagrama Back-End Calificación (Autores).

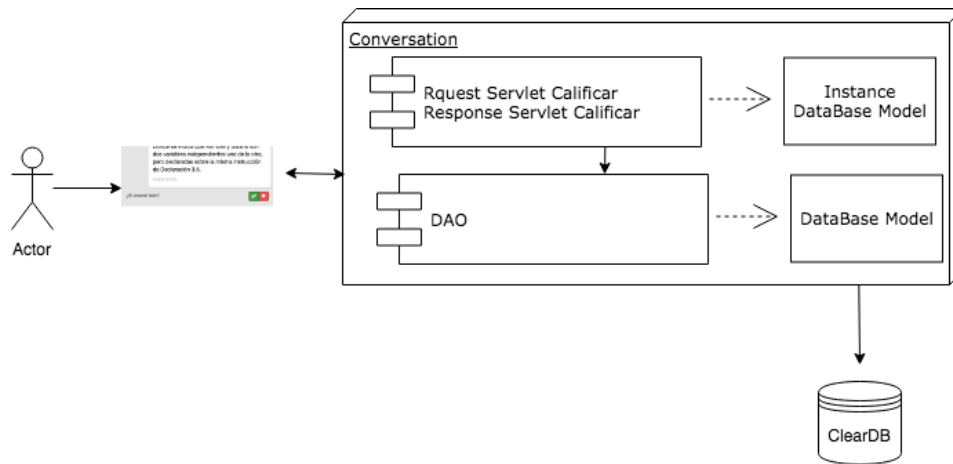
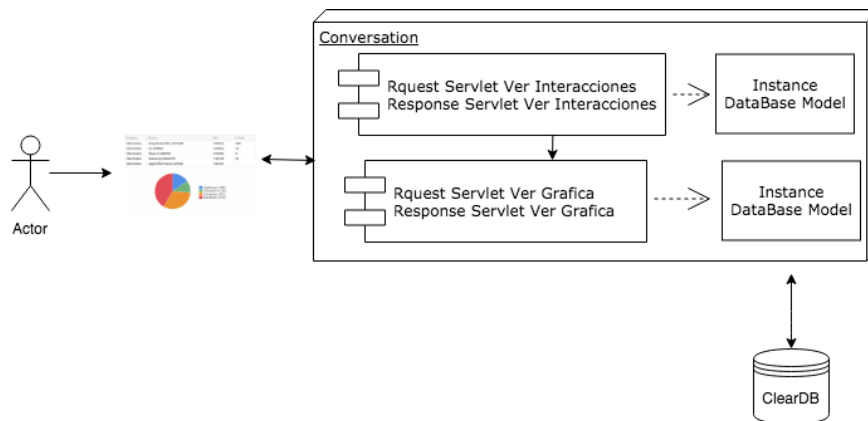


Figura 34 Diagrama Back-End Visualización (Autores).



7. ACEPTACIÓN DE USUARIOS

Según la encuesta realizada a 26 Ingenieros de sistemas acerca de la implementación del prototipo desarrollado (Arquímedes) se pudo determinar.

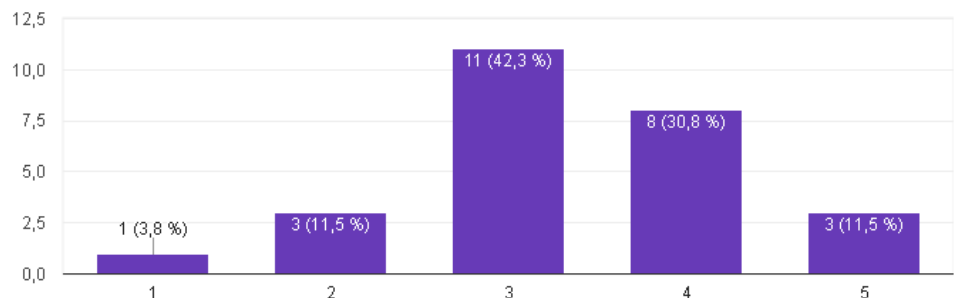
- La información brindada en los cursos virtuales en la mayoría de las veces no es lo suficientemente clara con un 42% de los estudiantes que seleccionaron 3, siendo uno que nunca es clara y 5 siempre es clara, lo anterior implica que estos estudiantes deben buscar apoyo en otras herramientas para así poder profundizar en los temas desarrollados o aclarar las dudas sobre estos.

Figura 35. Claridad en la información de los cursos virtuales.

¿La información brindada en los cursos virtuales referente a los temas a desarrollar es clara?



26 respuestas



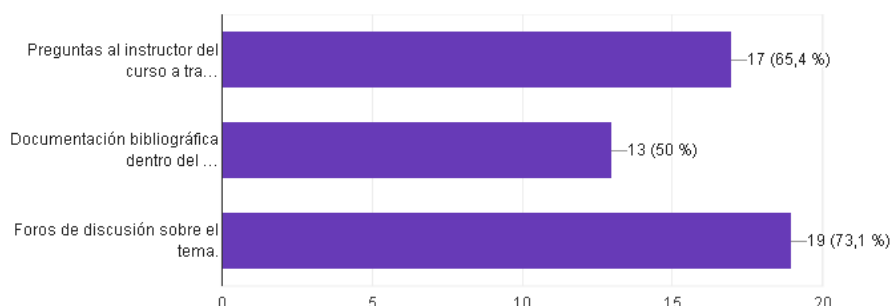
- La comunicación con los profesores es una de las opciones más utilizadas por los estudiantes para profundizar o aclarar los temas de las aulas virtuales, con un 65,4% de los estudiantes que usan esta opción en apoyo con los foros de discusión (73,1%) de los estudiantes.

Figura 36. Opciones utilizadas por los estudiantes para aclarar temas de las aulas virtuales.

Selecciona las opciones que has utilizado para complementar la información dentro del aula virtual.



26 respuestas

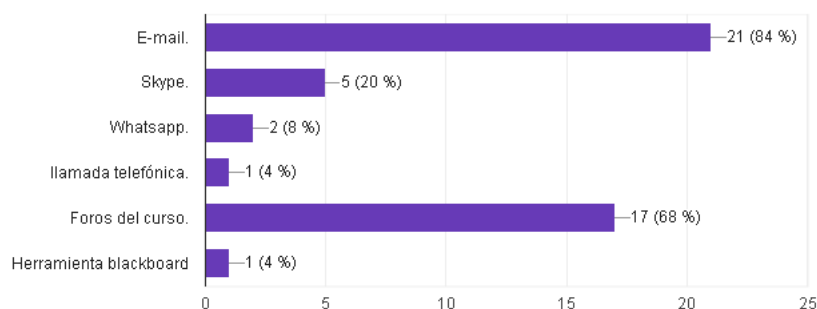


- La forma más común de realizar la comunicación con los docentes para aclarar las dudas es el correo electrónico seleccionado por 21 estudiantes (84%), seguida de los foros de discusión seleccionada por 17 estudiantes (68%).

Figura 37. Medios de comunicación más usados para contactar al docente.

¿Si ha intentado resolver dudas sobre el curso virtual a través de la comunicación con el instructor, que medios usa comúnmente?

25 respuestas



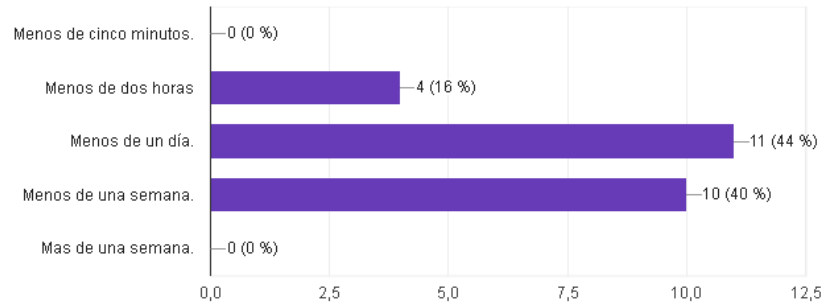
- Teniendo en cuenta que la forma más usada para contactar al docente es a través del e-mail, los estudiantes obtienen un rango de respuesta que llega a estar entre un día y una semana, nunca de forma instantánea; con un 46% de los estudiantes que reciben

respuestas en menos de un día y u 40% en menos de una semana.

Figura 38. Tiempo de respuesta de los instructores (Autores).

En promedio la respuesta por el instructor es dada en:

25 respuestas

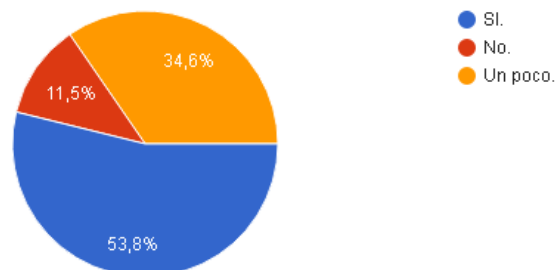


- Los estudiantes que interactuaron con Arquímedes consideraron que la información dada por este es relevante (14 estudiantes - 53,8%), un poco relevante (9 estudiantes - 34,6%) y no relevante con un (3 estudiantes - 11,5%).

Figura 39. Relevancia de la información otorgada por Arquimides (Autores)

Las respuestas dadas por Arquímedes aportaron información relevante?

26 respuestas

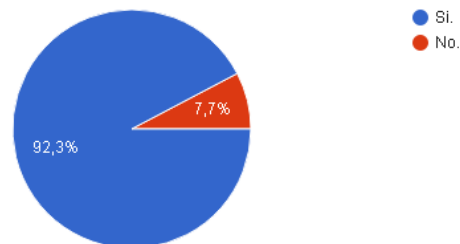


- También se encontró que el tiempo de respuesta de Arquímedes para 24 estudiantes (92,3%) es corto, y tan solo 2 estudiantes sintieron que no (7,7%).

Figura 40. Tiempo de respuesta de Arquímedes. (Autores)

Arquímedes respondió en corto tiempo?

26 respuestas

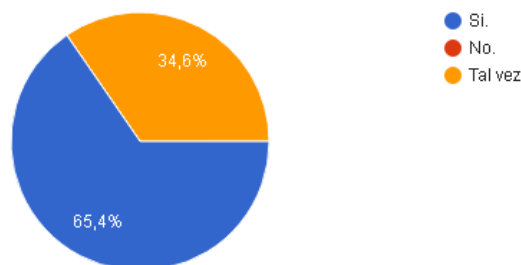


- Adicionalmente 17 estudiantes (65,4) expresaron que Arquímedes les parece útil como canal de comunicación adicional a las aulas virtuales y 9 estudiantes (34,6%) que este tal vez puede ser útil.

Figura 41. Utilidad de Arquímedes en las aulas virtuales (Autores)

Te parece Arquímedes útil como canal adicional para responder dudas en las aulas virtuales teniendo en cuenta que este mejorará sus respuestas en el tiempo?

26 respuestas

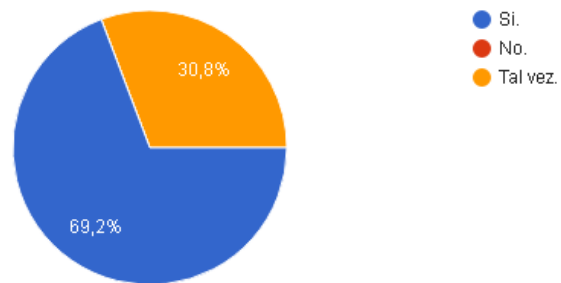


- Por ultimo 18 estudiantes referentes al 69,2% expresaron que les gustaría tener una herramienta como Arquimides en las diferentes aulas virtuales a las que acceden, y 8 estudiantes referentes al 30,8% que tal vez les gustaría tenerla en sus aulas virtuales.

Figura 42. Interés de Arquimides en las aulas virtuales (Autores)

Te gustaría tener una herramienta como Arquimides en las aulas virtuales a las que accedas?

26 respuestas



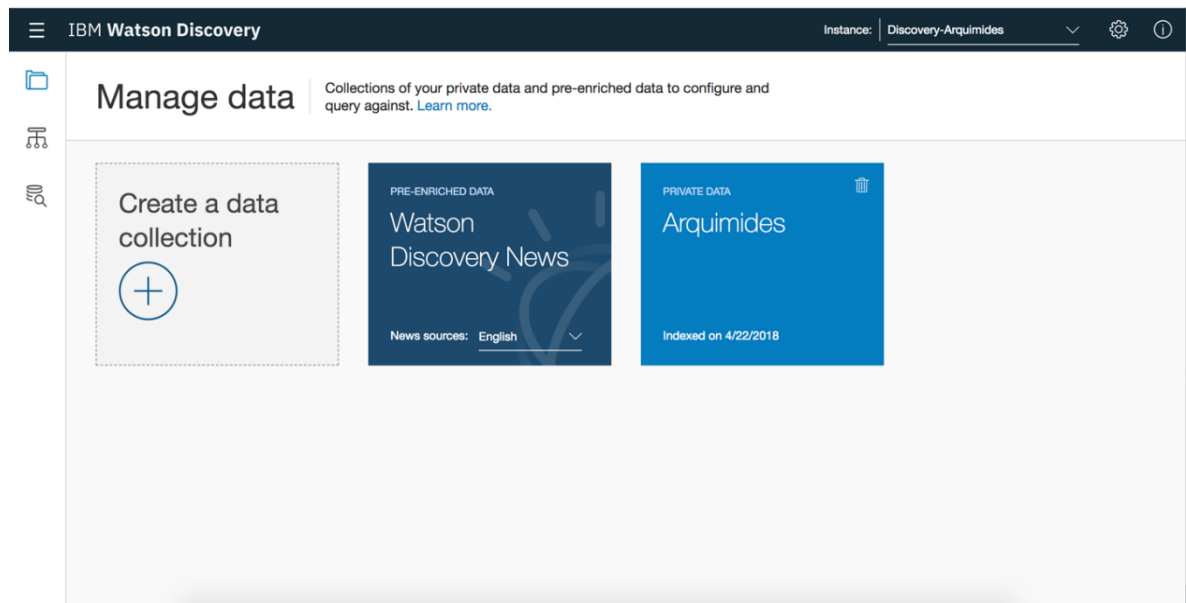
8. REQUERIMIENTO DE VALOR AGREGADO

Este apartado tiene como objetivo mostrar la metodología usada para obtener las respuestas de la herramienta “Watson Discovery Service”, un buscador cognitivo y analizador de contenido para identificar patrones sobre información no estructurada.

8.1. WATSON DISCOVERY SERVICE

Para utilizar este servicio es necesario tener una cuenta IBM Cloud, crear una instancia por medio del catálogo y lanzar la herramienta desde la instancia.

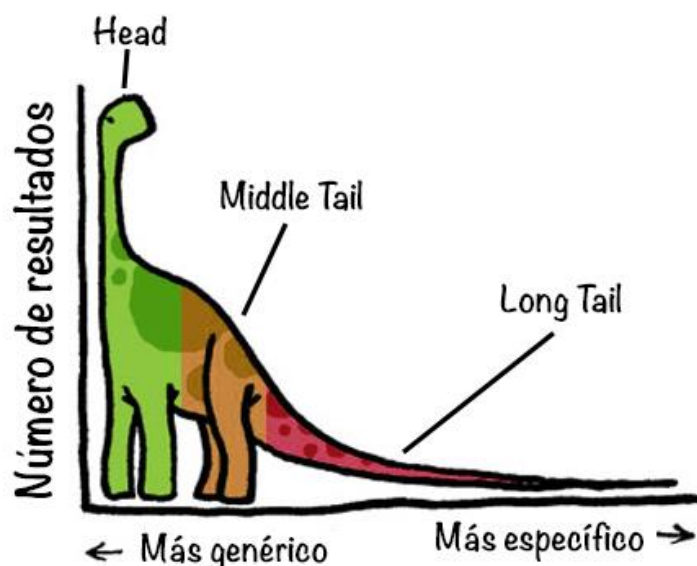
Figura 46. Página principal Watson Discovery (Autores).



El servicio brinda una interfaz gráfica para agregar, configurar y consultar contenido indexado, unificando estructurado y no estructurado con contenido pre enriquecido recibiendo HTML, DOC o PDF a JSON y consultándolas en lenguaje natural. Dentro de colecciones.

En el caso del proyecto, las preguntas de los estudiantes pueden ser muy variadas y específicas, por esto el servicio es utilizado para responder las preguntas dentro del “Long Tail”, preguntas menos frecuentes, pero de gran cantidad de las cual obtenemos la respuesta de documentos.

Figura 47. Short y Long Tail.⁶¹



Se crea una colección de datos con la configuración por defecto y el idioma español, abrirá automáticamente la pestaña de manejo de datos, en esta brindara información de los documentos subidos y la data enriquecida creada.

8.1.1 CONFIGURACIÓN DE DOCUMENTOS

La primera tarea es crear una configuración para los documentos acorde a la información que se requiere, para esto se crea una nueva configuración clicando el botón “switch “y a continuación seleccionando “Create a new configuration”.

⁶¹ **40deFiebre** ¿Qué es el Long Tail?. [En línea][Citado el. 17 04 2017.]<https://www.40defiebre.com/que-es/long-tail/>.

La pestaña “Convert” hace relación en las capacidades de conversión de Watson Discovery. Todos los documentos al final necesitan ser convertidos en formato JSON, para esto se selecciona PDF dentro de Convert y agrega los field que se necesiten definiendo el tamaño del texto y los estilos.

Para visualizar como la conversión está siendo realizada agrega un documento de muestra en la parte derecha de la pantalla.

Figura 48. Conversión en Watson Discovery (Autores)..

IBM Watson Discovery

Instance: Discovery-Arquimides

Convert text from your PDF documents into HTML

Set up rules for how you want your documents to be converted. [Learn more.](#)

Specify font properties

> H1 | assign when font is:

< H2 | assign when font is:

Font size range

18 to 24

Font style

☒ regular ☐ italic ☐ bold ☐ bold italic ☐ any style

Font name

e.g. Times New Roman

> H2 | assign when font is:

Upload sample documents

You can preview your configuration by using sample documents that are representative of your data collection. [Learn more.](#)

Drag and drop up to 10 documents or browse from computer

Sample documents expire after 24 hours

File format: PDF, WORD, HTML, and JSON

Maximum file size: 1MB

2D-Variables.pdf

m1t21.pdf

Last saved on: 4/30/2018 5:41:40 pm EDT

Apply & Save Close

A continuación, selecciona la opción HTML, esta configuración excluye o mantiene etiquetas, también puedes seleccionar XPath(s) específicas para mostrar, incluir o excluir contenido.

El último paso de conversión del documento es JSON, en donde según las etiquetas incluidas podemos combinar, renombrar o eliminar variables del objeto JSON final

8.1.2 ENRIQUECER DOCUMENTOS

Cuando se tienen definidos los campos retornados, se seleccionan los datos para enriquecer, enriquecer es la utilización de NLP para crear meta data del texto utilizándola para las búsquedas en lenguaje natural.


Agrega la variable del JSON donde se encuentra la información y a continuación selecciona “Add enrichments”, selecciona las que sean necesarias para la respuesta final de la herramienta.







Figura 49. Configuración de texto enriquecido (Autores).

Convert **Enrich** **Normalize**

Enrich your data with additional Watson insights

Set up rules for which fields you want to apply enrichments to. [Learn more.](#)

Add a field to enrich Choose a field 

Fields to be enriched	Enrichments
text	<div>keywords  concepts  semantic_roles </div> <div>entities  relations  elements </div> <div>+ Add enrichments</div>

• Last saved on: 5/1/2018 12:33:12 pm EDT

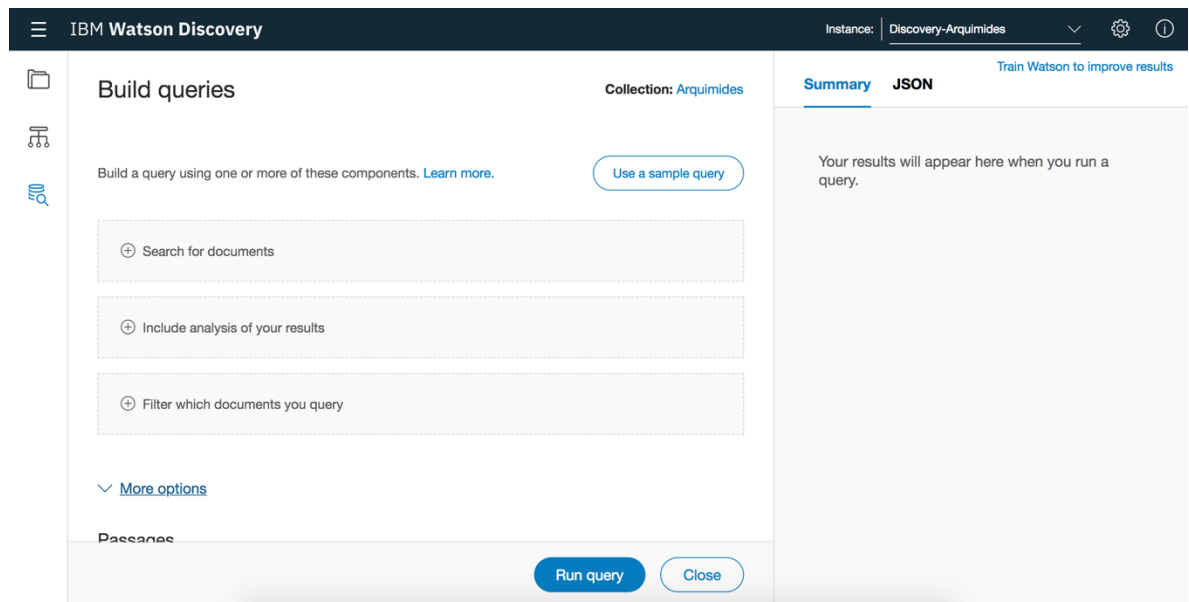
Apply & Save **Close**

Es importante tener en cuenta que cada texto enriquecido sube o baja la relevancia de las búsquedas, esto significa que agregar campos que no son de utilidad pueden afectar los resultados de la búsqueda.

8.1.3 ENTRENAR LA HERRAMIENTA PARA MEJORAR LOS RESULTADOS

La herramienta brinda la opción de mejorar las respuestas dadas por medio de entrenamiento, en el cual consiste en clasificar las respuestas dadas. Para esto selecciona Build Querys en la parte izquierda de la herramienta y a continuación Train Watson to improve results

Figura 50. Entrenamiento de la herramienta (Autores).



Agrega la query que se necesita mejorar y a selecciona rate results, selecciona relevant para los elementos que responden la pregunta e irrelevant para los elementos que no.

Figura 51. Clasificar respuestas (Autores)

Build queries >

Collection: Arquimides

Watson will learn which are the best results for your queries after you've rated enough.

Add more queries

Rate more results

Add more variety to your ratings

Fundamentos de Programación

View document

* ...

5.5. Manejo de Variables

El objetivo de las variables es permitir manejar cálculos parciales en el interior de un método. Las variables se deben declarar ...

... Las variables se deben declarar (darles un nombre y un tipo) antes de ser utilizadas y siguen la misma convención de nombres de los atributos. Las variables se crean en el momento en el que se declaran y se destruyen automáticamente al llegar al final del método que las contiene. ...

Show less

Relevant

Not relevant

Libro fundamentos de programacion

View document

* ... También que cada una de ellas recibe un nombre simbólico con el cuál pueda ser usado. El formato para la declaración de variables es el siguiente: Por ejemplo: Otra forma de declarar variables es también: Donde se indica que Nombre y Salario son dos variables independientes una de la otra, pero declaradas sobre la misma instrucción de Declaración 3.5. ...

... Ese fragmento del código indica que la variable Número1 se va a declarar. (5). Ese otro fragmento de código indica también que la variable Número2 se va a declarar. (6). La palabra Fin indica que se terminó un Pseudocódigo. Es importante recordar los tipos de datos para la declaración de variables; en la sección 3.3. ...

* ... También que cada una de ellas recibe un nombre simbólico con el cuál pueda ser usado. El formato para la declaración de variables es el siguiente: Por ejemplo: Otra forma

Microsoft Word - Portada.docx

View document

* ... En lugar de declarar N variables...

Página 375 Fundamentos de la programación: Tipos e instrucciones

125.40

76.95

328.80

* ... tipo de dato para varios elementos: Notas de los estudiantes de una clase Ventas de cada día de la semana Temperaturas de cada día del mes. En

Show more

Relevant

Not relevant

Cuando se consuma el servicio de la herramienta, este tendrá la configuración seleccionada en la herramienta visual y la aplicación podrá utilizar la información para determinar la respuesta al usuario.

9. RECOMENDACIONES

9.1. INTEGRACION CON LA PLATAFORMA E-LEARNING

El proyecto tiene como finalidad mostrar la utilización de tecnología cognitiva para resolver la comunicación síncrona en las aulas virtuales, esta herramienta puede abarcar más información al integrarse con la plataforma e-Learning que se encuentra, para guardar así estadísticas de uso, de calificación por curso, profesor y alumno, utilizar la capacidad de búsqueda cognitiva para herramientas de la plataforma fuera del chat.

9.2. ENTRADA/SALIDA POR VOZ

Ampliar las entradas salidas de la herramienta voz, para ampliar la accesibilidad de la herramienta y canales de utilización, en donde el canal puede variar consumiendo la misma instancia de la herramienta.

9.3. PREGUNTAS FRECUENTES

Agregar información de la clase. horarios, créditos, tutorías y prerrequisitos para preguntas del Short Tail que son menos complicadas de encontrar pero que son más frecuentes en los usuarios y orquestar la solución con las dos fuentes de información.

10.CONCLUSIONES

10.1. APORTACIONES DE LA TESIS

Se pudo implementar un canal de comunicación sincrónico el cual se logró a través de la herramienta Watson Discovery parametrizándola y refinándola para convertir y normalizar datos no estructurados por medio de preguntas en lenguaje natural, permitiendo integrar esta como una aplicación de software en forma de chat basada en una arquitectura orientada a servicios que permitió la interacción del usuario con la herramienta, siendo un factor diferencial con respecto a las soluciones investigadas que informan sobre la estructura de las asignaturas o servicios universitarios.

Se utilizó una herramienta de Análisis de data no estructurada como fuente de información de datos a diferencia de la implementación convencional para preguntas frecuentes que usan la metodología de set de datos con ejemplos de las formas de pregunta para clasificar la respuesta, utilizar algoritmos de entendimiento de máquina para la clasificación de preguntas fuera del set y respuestas por cada una de estas investigadas en antecedentes.

Se demostró que la aplicación es integrable en cualquier aula virtual además de aportar interés y valor en la forma que los estudiantes buscan consolidar sus conocimientos en las aulas virtuales ya que les brindo un espacio más eficaz para la solución de sus dudas.

BIBLIOGRAFIA

40defiebre ¿Qué es el Long Tail? [En línea]. - 17 de 04 de 2018. - <https://www.40defiebre.com/que-es/long-tail/>.

Accenture Centro de Alto rendimiento de Arquitectura Orientada a Servicios (SOA). [En línea]. - 2008. - 04 de 29 de 2018. - http://www.kybele.etsii.urjc.es/docencia/IS_LADE/2012-2013/Material/CAR%20Accenture%20-%20SOA%5B1%5D.pdf.

Corporation Microsoft La Arquitectura Orientada a Servicios (SOA) de Microsoft. [En línea]. - 12 de 2006. - 04 de 29 de 2018. - http://download.microsoft.com/download/c/2/c/c2ce8a3a-b4df-4a12-ba18-7e050aef3364/070717-real_world_soa.pdf.

ENCAMINA Revoluciona tu centro educativo con EduarBot [En línea]. - 17 de 03 de 2018. - <https://www.encamina.com/eduarbot/>.

FERNANDEZ-PAMPILLON. Ana Las plataformas e-learning para la enseñanza y el aprendizaje universitario en Internet. [En línea]. - 02 de 17 de 2018. - http://eprints.ucm.es/10682/1/capituloE_learning.pdf.

FONSECA Jorge, PAEZ Edwin y CORREDOR Emma Revista Ciencia y Agricultura Vol.13: Evaluación del uso de herramientas sincrónicas y asincrónicas en procesos de formación de las ciencias agropecuarias. [En línea]. - 20 de 04 de 2018. - <https://dialnet.unirioja.es/descarga/articulo/5560523.pdf>.

García Óscar Unit4 presenta su primer chatbot para el estudiante de Educación Superior [En línea]. - 25 de 10 de 2016. - 17 de 03 de 2017. - <https://www.unit4.com/es/acerca-de/noticias/2016/10/unit4-presenta-su-primer-chatbot-para-el-estudiante-de-educacion-superior>.

Google LLC Dialogflow Docs [En línea]. - 21 de 02 de 2018. - 17 de 03 de 2018. - <https://dialogflow.com/docs>.

IBM About Discovery [En línea]. - 02 de 02 de 2018. - 17 de 3 de 2018. - <https://console.bluemix.net/docs/services/discovery/index.html#about>.

IBM About Watson Assistant [En línea]. - 26 de 01 de 2018. - 17 de 03 de 2018. - <https://console.bluemix.net/docs/services/conversation/index.html#about>.

IBM Arquitectura orientada a servicios [En línea]. - 17 de 02 de 2018. - https://www.ibm.com/support/knowledgecenter/es/SSFPJS_8.5.6/com.ibm.wbpm.wid.main.doc/prodoverview/topics/csoa.html.

IBM Discovery pricing plans [En línea] // IBM Cloud. - 23 de 10 de 2018. - <https://console.bluemix.net/docs/services/discovery/pricing-details.html#discovery-pricing-plans>.

IBM Models for machine learning [En línea] // IBM developerWorks®. - 05 de 12 de 2017. - <https://www.ibm.com/developerworks/library/cc-models-machine-learning/index.html>.

IBM Pricing Calculator [En línea] // IBM Cloud. - 23 de 10 de 2018. - <https://console.bluemix.net/pricing/configure/runtime/javaHelloWorld>.

Izquierdo Robin NoSQL vs SQL: Principales diferencias y cuándo elegir cada una de ellas [En línea]. - 2015. - 19 de 09 de 2018. - <https://blog.pandorafms.org/es/nosql-vs-sql-diferencias-y-cuando-elegir-cada-una/>.

Kelly E. John Why cognitive systems? [En línea] // IBM Research. - IBM, 10 de Junio de 2016. - 3 de Abril de 2017. - <http://www.research.ibm.com/cognitive-computing/why-cognitive-systems.shtml#fbid=5oarVDsWcsr>.

LOPEZ . Bruno Introducción a la inteligencia Artificial. [En línea]. - 03 de 10 de 2018. - <http://www.itnuevolaredo.edu.mx/takeyas/Articulos/Inteligencia%20Artificial/ARTICULO%20Introduccion%20a%20la%20I>.

Malaeb Maher Recall and Precision at k for Recommender Systems [En línea] // Medium. - 13 de 8 de 2017. - 1 de 11 de 2018. - https://medium.com/@m_n_malaeb/recall-and-precision-at-k-for-recommender-systems-618483226c54.

MORIELLO Sergio Cerebro, Mente, Cuerpo y Entorno. [En línea]. - 24 de 02 de 2018.. - http://www.pensamientocomplejo.com.ar/docs/files/Moriello_Cerebro-mente-cuerpo-entorno.pdf.

Organización para la Cooperación y el Desarrollo Económicos (OCDE) Educación en Colombia: Educación superior en Colombia. [En línea]. - 2016. - 03 de 02 de 2017. - https://www.mineduacion.gov.co/1759/articles-356787_recurso_1.pdf.

Peña Daniel Herramientas SQL vs NoSQL [Informe]. - 2015.

PEREZ Cevallos Ciencia y tecnología al alcance de todos. [En línea]. - 2007. - 03 de 02 de 2017. - <http://www.bib.uia.mx/tesis/pdf/014866/014866.pdf>.

Real Academia Española Dato [En línea] // DEL Real Academia Española. - 10 de Junio de 2014. - 3 de Abril de 2017. - <http://dle.rae.es/?id=Bskzsq5|BsnXzV1>.

Real Academia Española usuario [En línea] // DLE Real Academia Española. - 13 de Junio de 2014. - 3 de Abril de 2017. - <http://dle.rae.es/?id=bBsqKPc>.

Rodríguez del Rey Lamí, L. E., Pérez Fleites, M. G., & Rodríguez del Rey Rodríguez, M. E. Las herramientas de comunicación sincrónica y asincrónica en la clase presencial. [Artículo] // Revista Conrado. - [s.l.] : Revista Conrado, 2016. - 12 (56), pp. 84-89..

SAS Natural Language Processing [En línea] // SAS. - https://www.sas.com/es_co/insights/analytics/what-is-natural-language-processing-nlp.html.

Scagnoli Norma El aula virtual: usos y elementos que la componen. [Informe] / Universidad de Illinois. - 2001.

SCHWABER Ken La Guía Definitiva de Scrum: Las Reglas del Juego. [En línea]. - 2016. - 17 de 03 de 2018. - <http://www.scrumguides.org/docs/scrumguide/v2016/2016-Scrum-Guide-Spanish.pdf>.

TRIGAS. Manuel Gestión de Proyectos Informáticos. [En línea]. - 17 de 03 de 2018. - <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/17885/1/mtrigasTFC0612memoria.pdf>.

TRIGAS. Manuel Gestión de Proyectos Informáticos. [En línea]. - 03 de 17 de 2018. - <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/17885/1/mtrigasTFC0612memoria.pdf>.

WIECHERS Rivero Lógica, versión icónica. Editorial humanismo y sentido. [En línea]. - México 2009.. - 24 de 02 de 2018. - http://cvonline.uaeh.edu.mx/Cursos/Bach_Virt/H0102/unidad_4/lec_r10315_el_razonamiento.pdf.

ANEXOS

ANEXO A. FORMATO DE ENCUESTA

Aulas virtuales

Esta encuesta tiene como finalidad buscar la percepción de la comunicación en las aulas virtuales.

¿Ha participado usted en cursos con aulas virtuales?

☐ Sí

☐ No

¿La información brindada en los cursos virtuales referente a los temas a desarrollar es clara?

	1	2	3	4	5	
Nunca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Siempre

Selecciona las opciones que has utilizado para complementar la información dentro del aula virtual.

☐ Preguntas al instructor del curso a través de medios de comunicación.

☐ Documentación bibliográfica dentro del curso virtual.

☐ Foros de discusión sobre el tema.

¿Si ha intentado resolver dudas sobre el curso virtual a través de la comunicación con el instructor, que medios usa comúnmente?

- ☐ E-mail.
- ☐ Skype.
- ☐ Whatsapp.
- ☐ llamada telefónica.
- ☐ Foros del curso.
- ☐ Otra...

En promedio la respuesta por el instructor es dada en:

- ☐ Menos de cinco minutos.
- ☐ Menos de dos horas
- ☐ Menos de un día.
- ☐ Menos de una semana.
- ☐ Mas de una semana.

Interacción Con Arquimides.

Arquimides es un prototipo en forma de chat que puede responder tus dudas a cerca de fundamentos de programación y aprender de las interacciones para mejorar sus respuestas.

Para interactuar con Arquimides puedes dirigirte a:

<http://arquimidesapp.mybluemix.net/>

Las respuestas dadas por Arquimides aportaron información relevante?

- ☐ Si.
- ☐ No.
- ☐ Un poco.

Arquimides respondió en corto tiempo?

☐ Si.

☐ No.

Te parece Arquimides útil como canal adicional para responder dudas en las aulas virtuales teniendo en cuenta que este mejorará sus respuestas en el tiempo?

☐ Si.

☐ No.

☐ Tal vez

Te gustaría tener una herramienta como Arquimides en las aulas virtuales a las que accedas?

☐ Si.

☐ No.

☐ Tal vez.

ANEXO B. PLAN DE PRUEBAS

- **Alcance**

Las pruebas serán realizadas desde la url donde se encuentra dispuesto la interfaz del chat y del Dashboard.

Se validará el despliegue del componente del chat, poder realizar la pregunta y que responda en máximo 10 segundos.

Se validará que la interacción pueda calificarse correctamente y que las calificaciones coincidan con los datos almacenados en la base de datos ClearDB.

Se validará que la interfaz del Dashboard de despliegue correctamente, que muestre la tabla de interacciones y la gráfica de calificaciones, con la información acorde a los datos almacenados en la base de datos ClearDB.

- **Fuera del alcance**

Validar informes diferentes descritos en el alcance.

Validar el apartado funcional del chat y Dashboard.

Validar las pruebas especializadas, tanto de seguridad como de performance, esto debido a ser un prototipo y estar en un ambiente limitado que no es el óptimo para producción.

- **Estrategia**

Se limpiará la base de datos para cada escenario de prueba que involucre almacenar en la base de datos.

Para las pruebas se generarán datos aleatorios por medio de script para probar el comportamiento del modelo de datos.

Se realizarán preguntas al chat dentro y fuera del alcance para probar el comportamiento del bot.

Se calificarán respuestas aleatoriamente dentro del chat para probar el comportamiento del bot.

Se revisará los indicadores del Dashboard tanto cargados con Script como por el chat para revisar la concordancia con los almacenados en la base de datos.

Los casos de prueba se alternarán el cargo desarrollador y certificador entre el equipo de trabajo para validar la funcionalidad imparcialmente.

- **Equipo de trabajo**

Gustavo Millan	Ivan Arce	Desarrollo Certificación	/
Omar Ortiz	David Sánchez	Desarrollo Certificación	/

- **Prerrequisitos**

Se debe de contar con la solución instanciada en el ambiente Cloud.

Contar con permisos de lectura y escritura sobre la base de datos ClearDB.

- **Supuestos**

Se cuenta con el ambiente de pruebas estable.

El ambiente de pruebas se encuentra configurado de forma correcta.

Los defectos reportados por el equipo de certificación serán atendidos por el equipo de desarrollo o el proveedor de la solución en los tiempos pactados.

Se contará con el apoyo del proveedor y el analista de desarrollo, para la solución de dudas o errores que se presenten en la ejecución de las pruebas.

- **Limitaciones**

No se cuenta con los recursos (humanos, físicos, entre otros).

Problema de acceso a los aplicativos que se requieren para la ejecución de la prueba.

Los servicios adicionales para la consulta no funcionen de manera correcta.

ANEXO C. Casos de prueba

Tabla 5. Caso CH1 (Autores)

Caso de prueba	10.1.1 Respuesta del chat		
Identificador caso de prueba	CH1		
Función probar	<i>El componente chat se despliegue y de respuesta cuando se interactúe con el</i>		
Objetivo	<i>El chat responde a las interacciones</i>		
Criterios de éxito	<i>El componente comenzara compreso en una barra en la parte inferior derecha</i> <i>El componente se desplegará cuando se seleccione el mas o cuando se seleccione el campo de texto</i> <i>El componente aceptara texto en el campo de texto, se pueda enviar correctamente al presionar “Enter”</i> <i>El componente muestra correctamente el mensaje enviado en el campo superior del campo de texto</i> <i>El chat responde a la interacción</i>		
Criterios de falla	<i>El chat no se muestra en la interfaz</i> <i>El chat no se despliega correctamente</i> <i>El chat no responde cuando el usuario envía una interacción</i>		
Precondiciones	<i>Descritas en el plan de pruebas</i>		
Perfil del usuario	<i>Funcional</i>		
Necesidades para el caso de prueba	<i>Descritas en el plan de pruebas</i>		
Autor	<i>Gustavo Ivan Arce Millan</i>		
Fecha de creación	07/04/2018		
Flujo del caso de prueba	No paso	Usuario del sistema	Sistema
	0	<i>Flujo correspondiente al caso de uso No. 001</i>	<i>Respuesta correspondiente al caso de uso No. 01</i>

Tabla 6. Caso CH1 (Autores)

Caso de prueba	10.1.2 Calificación de la interacción		
Identificador caso de prueba	CH2		
Función probar	La aplicación guarde en la base de datos al calificar en la conversación una interacción		
Objetivo	Garantizar que la calificación suministrada por el usuario sea guardada correctamente en la base de datos		
Criterios de éxito	<p>El componente muestra correctamente después de una interacción los botones de calificar ✓ y X</p> <p>Al seleccionar cualquiera de los dos botones, la información quede almacenada en la base de datos con los campos correspondiente a interacción, respuesta y calificación</p>		
Criterios de falla	<p>No se muestra correctamente los botones</p> <p>No se guarda correctamente la calificación</p>		
Precondiciones	Descritas en el plan de pruebas Ejecutado el Caso de uso 001		
Perfil del usuario	Técnico		
Necesidades para el caso de prueba	Descritas en el plan de pruebas		
Autor	Omar David Sánchez Ortiz		
Fecha de creación	05/04/2018		
Flujo del caso de prueba	No paso	Usuario del sistema	Sistema
	1		Se muestra los

			componentes de calificación
	2	Usuario califica la interacción	Se captura la pregunta, la respuesta y la calificación y se guarda en la tabla Questions
		Se verifica que la interacción este guardada en la base de datos mediante acceso a la instancia y consulta	

Tabla 7. Caso CH1 (Autores)

Caso de prueba	10.1.3 Calificación de la interacción
Identificador caso de prueba	CH3
Función probar	La aplicación guarde en la base de datos al calificar en la conversación una interacción
Objetivo	Garantizar que la calificación suministrada por el usuario sea guardada correctamente en la base de datos
Criterios de éxito	El componente muestra correctamente después de una interacción los botones de calificar ✓ y X Al seleccionar cualquiera de los dos botones, la información quede almacenada en la base de datos con los campos

	<i>correspondiente a interacción, respuesta y calificación</i>		
Criterios de falla	<i>No se muestra correctamente los botones</i> <i>No se guarda correctamente la calificación</i>		
Precondiciones	<i>Descritas en el plan de pruebas</i> <i>Ejecutado el Caso de uso 001</i>		
Perfil del usuario	<i>Técnico</i>		
Necesidades para el caso de prueba	<i>Descritas en el plan de pruebas</i>		
Autor	<i>Omar David Sánchez Ortiz</i>		
Fecha de creación	<i>05/04/2018</i>		
Flujo del caso de prueba	No paso	Usuario del sistema	Sistema
	1		<i>Se muestra los componentes de calificación</i>
	2	<i>Usuario no califica la interacción e interactúa</i>	<i>Se inhabilita los componentes de calificación y se procede con el flujo del caso de uso 001</i>
	3		<i>Se muestra los componentes de calificación</i>
	4		<i>Se captura la pregunta, la respuesta y la calificación de la última</i>

			<i>interacción y se guarda en la tabla Questions</i>
	5	<i>Se verifica que la interacción este guardada en la base de datos mediante acceso a la instancia y consulta</i>	

Tabla 8. Caso CH1 (Autores)

Caso de prueba	10.1.4 Visualización de graficas
Identificador caso de prueba	<i>CH4</i>
Función probar	<i>La interfaz se muestra correctamente en la página del Dashboard</i>
Objetivo	<i>Garantizar que los datos suministrados se muestren correctamente y sean coherentes a los almacenados en la base de datos</i>
Criterios de éxito	<i>El componente muestra correctamente al ingresar a la página La tabla de interacciones muestra la información acorde a la base de datos</i>
Criterios de falla	<i>No se muestra correctamente la interfaz Los datos no son acordes a los almacenados en la base de datos</i>
Precondiciones	<i>Descritas en el plan de pruebas Caso de uso 001 Caso de uso 002</i>
Perfil del usuario	<i>Técnico</i>
Necesidades para el caso de prueba	<i>Descritas en el plan de pruebas</i>
Autor	<i>Gustavo Ivan Arce Millan</i>

Fecha de creación	07/04/2018		
Flujo del caso de prueba	No paso	Usuario del sistema	Sistema
	1	Se accede a la url <i>context_root/Dahshboard.html</i>	Se muestra la tabla y la gráfica de calificaciones
	2	Se verifica que la información suministrada por la página sea coherente con la base de datos mediante acceso a la instancia y consulta	

ANEXO D. SQL vs NoSQL

Tabla 9. SQL vs NoSQL (Autores)

	SQL	NO SQL
1	Mayor soporte y add ons	Problema de compatibilidad SQL
2	Transaccional	
3	Integridad	Adaptabilidad
4	Menor escalabilidad	Escalabilidad
5		Esquemas
6	Escalabilidad vertical	Escalabilidad horizontal
7		Rapidez de consulta

1. Por su tiempo en el mercado las bases de datos no relacionales tienen mayor soporte y gestión de implementaciones.
Las bases de datos no SQL tienen problemas de compatibilidad entre instrucciones SQL, ya que estas utilizan sus propias características en el lenguaje de consulta y no son 100% compatibles con el SQL de las bases de datos relacionales. El soporte a problemas con las queries de trabajo en una base de datos NoSQL es más complicado, esto debido también a la falta de estandarización.
2. Esto hace referencia a que en las bases de datos relacionales van enfocadas hacia operaciones que se deben terminar completamente o se hace un rollback desde el punto en el que haya quedado impidiendo la finalización de esta.
3. Los datos deben cumplir con las condiciones como tipo de datos y compatibilidad.
Las bases de datos no SQL permiten el manejo de datos más abiertos y flexibles
4. Carácter descentralizado, soporta estructuras distribuidas.
5. En las bases de datos no SQL Se pueden realizar cambio de esquemas sin detener las bases de datos.
6. Escalabilidad horizontal. son capaces de crecer en número de máquinas, en lugar de tener que residir en grandes máquinas.
Escalabilidad vertical.
7. En las bases de datos SQL las consultas se pueden optimizar a través de la indexación, en cambio por la estructura sencillas de las bases de datos no SQL (clave - valor) se logra una mejor y más fácil optimización de las consultas para grandes cantidades de datos.

- **Cuando usar SQL**

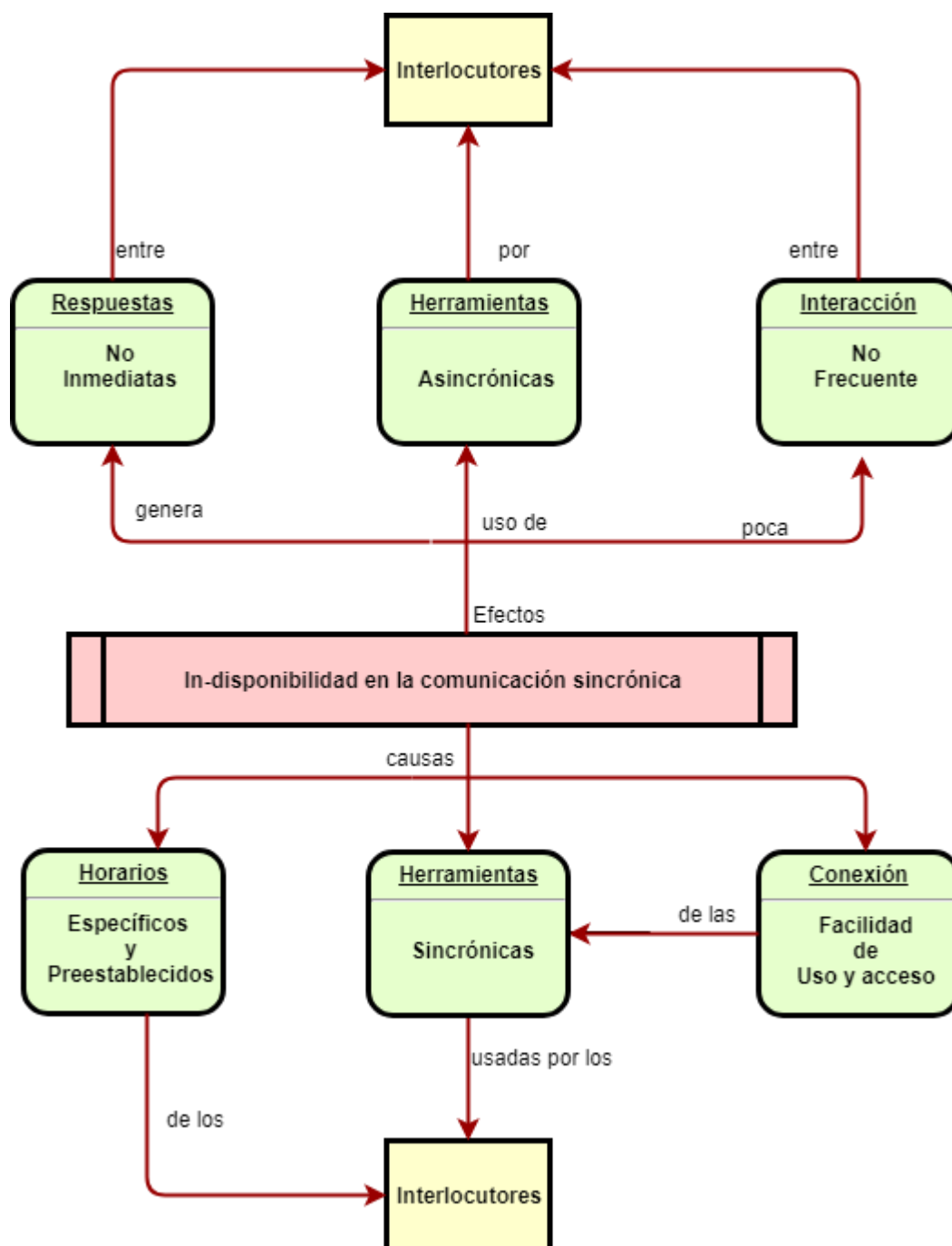
- Cuando los datos deben ser consistentes sin dar posibilidad al error.
- Para Business Intelligence, ya que estas se pueden enfocar en temas como minería de datos, bodegas de datos, minería de textos ya que SQL facilita el consumo, análisis y detección de patrones.

- **Cuando usar no SQL**

- Cuando nuestro presupuesto no se puede permitir grandes máquinas y debe destinarse a máquinas de menor rendimiento.
- Cuando los datos son no estructurados o poco uniformes como datos de internet.
- Análisis de grandes cantidades de datos en modo lectura, Big Data.
- Captura y procesamiento de eventos.

ANEXO E. DIAGRAMA INDISPONIBILIDAD HERRAMIENTAS SÍNCRONAS

Ilustración 43. Indisponibilidad (Autores)



ANEXO F. COSTOS Y LICENCIAMIENTO

El proyecto esta diseñado para utilizar y ser alojado en instancias Cloud, en este apartado se mostrará los recursos usados para el desarrollo que varían en ambientes productivos. Los productos Watson IBM solo son ofertados en su Infraestructura y son usables desde se tenga acceso a ellos.

- **Alojamiento de la aplicación**

En los ambientes de desarrollo la aplicación web es desplegada en la infraestructura nube Cloud Foundry en IBM Cloud, para propósitos de desarrollo se puede utilizar 256MB de RAM gratis que, si no se hacen modificaciones en la instancia, se desactivara a los 15 días.

Tabla 10. Comparación cuentas IBM Cloud

	Lite	Pay-As-You-Go	Subscription
Access to free Cloud Foundry memory	256 MB	512 MB	512 MB
Access to Lite service plans External link icon	✓	✓	✓
Access to all free plans		✓	✓
Access to the full IBM Cloud catalog		✓	✓
Access to multiple Cloud Foundry regions		✓	✓
No time restrictions	✓	✓	✓
Guaranteed zero cost	✓		
Discounted pricing			✓
Best for learning or building proof of concepts	✓	✓	
Fit for production use cases		✓	✓

El costo del alojamiento de la aplicación en producción en IBM Cloud varia de la memoria de la instancia, siendo gratis hasta los 512MB y un promedio de 5.25 USD por 128 MB después de los primero 512MB (IBM). La aplicación es posible desplegarla en un servidor de aplicaciones Apache o Liberty Profile.

- **Servicio Watson Discovery (Malaeb, 2017)**

Dentro de las herramientas utilizadas para la aplicación esta el servicio Watson alojado en IBM llamado Discovery, este servicio ofrece una versión Lite que se utiliza para ambientes de desarrollo y pruebas de manera gratuita, para ambientes productivos se lista muestran en la siguiente tabla (IBM).

Tabla 11. Tipos de plan en Watson Discovery

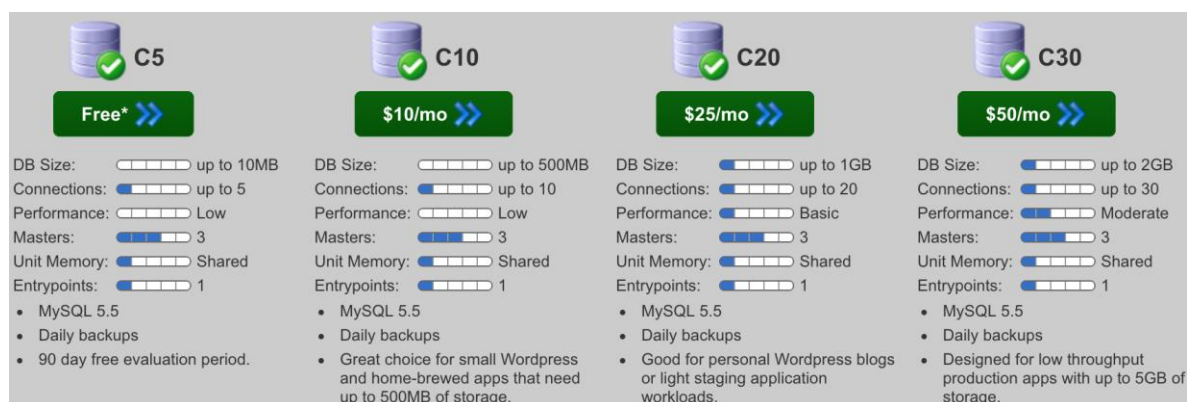
Nombre	Tamaño limite de la documentaci3n	Numero de documentos	Precio
Lite	50 MB	1,000 documentos por mes	Free
X-Small*****	40 GB	Up to 50,000 documentos por mes	Comenzando en \$500 USD por mes
Small	160 GB	Up to 1M documentos por mes	Comenzando en \$1,500 USD por mes
Medium-Small	320 GB	Up to 2M documentos por mes	Comenzando en \$3,000 USD por mes
Medium	640 GB	Up to 4M documentos por mes	Comenzando en \$5,000 USD por mes
Medium-Large	1.2 TB	Up to 8M documentos por mes	Comenzando en \$10,000 USD por mes
Large	2.4 TB	Up to 16M documentos por mes	Comenzando en \$15,000 USD por mes
X-Large	4 TB	Up to 32M documentos por mes	Comenzando en \$20,000 USD por mes
XX-Large	5.5 TB	Up to 64M documentos por mes	Comenzando en \$35,000 USD por mes
XXX-Large	12 TB	Up to 100M documentos por mes	Comenzando en \$45,000 USD por mes

Según la estimación de documentación es recomendable la instancia Lite y migrar a X-Small cuando se necesite mas capacidad.

ANEXO G. Base de datos

Para guardar las interacciones y hacer seguimiento del mejoramiento de la herramienta se necesita una base de datos que persista la aplicación. Para propósitos de desarrollo se ha utilizado una base de datos relacional ClearDB basada en MYSQL, alojada en Google Cloud Plataform gratuitamente durante 90 días.

Figura 44. Planes de ClearDB



Para los ambientes productivos se recomienda la instancia C20 de ClearDB en la nube Google Cloud Plataform, también es posible agregar el esquema a una base de datos local si se dispone de una.

El costo de la solución en producción puede variar entre 75 a 525 USD en infraestructura Cloud dependiendo del trafico y tamaño/alcance de la solución.

ANEXO H. PRUEBAS DE LA HERRAMIENTA

Uno de los factores mas importantes a la hora de implementar herramientas basadas en Machine Learning es tener en cuenta que los resultados son basados en estadísticas y los resultados no son binarios. Para buscar un buen rendimiento en ambientes productivos existen métodos para testear la herramienta y generar medidas que ayuden a tomar decisiones sobre si la herramienta necesita refinamiento o esta lista para desplegarse.

Para el caso de la aplicación Arquímedes, que utiliza un motor de búsqueda basado en herramienta cognitiva Watson Discovery, se utilizan métodos que aplican para aprendizaje de clasificación y para modelos de entrenamiento supervisados, existen tres tipos de métodos mas usados. Precision@k, Mean Average Precision y Mean Reciprocal Rank, el aplicado en el proyecto es el Precision@K, debido a que la complejidad de la búsqueda se basa directamente en las variables generadas con el enriquecimiento del texto, que podemos controlar.

- **Precisión y Recall de K**

Son métricas binarias usadas para evaluar salidas binarias, por lo que tenemos que traducir el problema de un numero a la salida valida.

El primer paso para esto es definir un numero limite para que una recomendación sea relevante o irrelevante (1 o 0), en este caso se definió como relevante los resultados por encima de 0.6, este valor puede ser definido teniendo en cuenta historial de clasificación de usuarios en canales similares.

El siguiente paso es definir la K, esto basado en que lo que nos interesa es revisar el top de resultados que se le da al usuario por lo que es mas recomendable calcular los resultados de la Precisión y Recall del primer N ítem, en vez de todos los posibles resultados.

Al tener claro el numero que definirá los resultados relevantes y el computo del N ítem mas alto, es necesario definir la Relevancia y lo Recomendado. los ítems relevantes son aquellos que son iguales o superan el limite establecido y que conocemos previamente como relevantes, lo Recomendado son aquellos que son predichos igual o por encima del limite establecido.

- **Lo Recomendado con lo relevante**

La precisión de K es la proporción de ítems recomendados que son relevantes donde k es el limite de resultados, aplicado se define como.

$\text{Precision@k} = (\# \text{ recomendados que son relevantes} / (\# \text{ de ítems recomendados}))$

El Recall de K es la proporción de ítems relevantes encontrados en el top de recomendados, aplicado se define como.

$\text{Recall@k} = (\# \text{ recomendados que son relevantes}) / (\# \text{ total de ítems relevantes})$

Para poder poner en practica estas formulas, es necesario crear un set de datos de entrenamiento. preguntas que se tengan el valor “actual”, la relevancia de las respuestas a esta pregunta, para que al momento de utilizar el algoritmo se pueda comparar lo actual (llamado al numero relevante) con la salida del algoritmo (llamado al numero lo recomendado). Estas pruebas tienen que hacerse con muestras aleatorias de información nombrada en la documentación agregada.

- Se recomienda generar un stack de mínimo 10 ejemplos de un tópico seleccionado (contenido de la documentación)
- Los ejemplos deben probarse en la herramienta y, de las respuestas datas clasificar cual de estas es relevante
- Con los mismos resultados, se compara cada elemento las respuestas clasificadas manualmente con las respuestas del algoritmo, las respuestas clasificadas como relevantes manualmente son el numero relevante, y las del algoritmo las recomendadas
- Se aplican las formulas con los datos obtenidos

Estos dos datos nos dicen cual es el porcentaje de elementos relevantes de nuestra herramienta (precisión) y el porcentaje de elementos que mostramos relevantes de los relevantes existentes (Recall).

Estas pruebas son aplicadas iterativamente hasta buscar un porcentaje entre mínimo el 78% en estos indicadores.

- **Resultados**

Tabla 12. Resultados del Precision@K

input	relevantes	recomendados	recomendados relevantes
declarar variable	0.5970597	0.58893573	
que es un constructor	1.2515876	1.2551719	1
paradigmas de programación	1.5845512	1.1145634	1
estructuras de programación	0.8758188 2	0.723756	1
listas de listas	0.7028723 1	0.9854727	1
abstracción	0.8787130 5	0.252868146	
instanciación de objetos	0.4966922 1	0.584609862	
definición de clases	0.6999819 7	0.503618185	
usos de interfaces	0.2997107 3	0.85104847	
java swing	0.7384079	0.030574218	

**Numero con limite de
0.68**

7

5

4

Precision. $4/5 = 0.8$

Recall. $4/7 = 0.57$

Con estos datos el sistema puede ser mejorado para tener una mayor relevancia sobre ítems que son aun mas relevantes que el top, pero tiene una relevancia aceptable para dar al usuario.

Es de aclarar que esta medida debe ser tomada constantemente debido a la recomendación dada de continuar entrenando la herramienta en pasos posteriores al despliegue en producción.